Science Library

JUL 2 4 1945

Psychometrika

A JOURNAL DEVOTED TO THE DEVELOPMENT OF PSYCHOLOGY AS A QUANTITATIVE RATIONAL SCIENCE

THE PSYCHOMETRIC SOCIETY · ORGANIZED IN 1935

VOLUME 10
NUMBER 2
JUNE
1945

PSYCHOMETRIKA, the official journal of the Psychometric Society, is devoted to the development of psychology as a quantitative rational science. Issued four times a year, on March 15, June 15, September 15, and December 15.

JUNE 1945, VOLUME 10, NUMBER 2

Printed for the Psychometric Society at 23 West Colorado Avenue, Colorado Springs, Colorado. Entered as second class matter, September 17, 1940, at the Post Office of Colorado Springs, Colorado, under the act of March 3, 1879. Editorial Office, College Entrance Examination Board, Princeton, New Jersey.

Subscription Price:

To non-members, the subscription price is $5.00 per volume of four issues. Members of the Psychometric Society pay annual dues of $5.00, of which $4.50 is in payment of a subscription to *Psychometrika*. Student members of the Psychometric Society pay annual dues of $3.00, of which $2.70 is in payment for the journal.

The subscription price to libraries and other institutions is $10.00 per year; this price includes one extra copy of each issue.

Applications for membership and student membership in the Psychometric Society should be sent to

THELMA GWINN THURSTONE
Chairman of the Membership Committee
University of Chicago
Chicago, Illinois

Membership dues are payable to

IRVING D. LORGE
Treasurer of the Psychometric Society
425 West 123rd Street
New York, New York

Individual and library subscriptions are payable to, and change of address should be sent to

IRVING D. LORGE
Treasurer of the Psychometric Corporation
425 West 123rd Street
New York, New York

Articles on the following subjects are published in *Psychometrika*:

(1) the development of quantitative rationale for the solution of psychological problems;
(2) general theoretical articles on quantitative methodology in the social and biological sciences;
(3) new mathematical and statistical techniques for the evaluation of psychological data;
(4) aids in the application of statistical techniques, such as monographs, tables, work-sheet layouts, forms, and apparatus;
(5) critiques or reviews of significant studies involving the use of quantitative techniques.

(*Continued on the back inside cover page*)

Psychometrika

CONTENTS

A MULTIPLE GROUP METHOD OF FACTORING THE CORRELATION MATRIX - - - - - - - - 73
L. L. THURSTONE

THE RELATION OF ITEM DIFFICULTY AND INTER-ITEM CORRELATION TO TEST VARIANCE AND RELIABILITY - - - - - - - - - - - - 79
HAROLD GULLIKSEN

FACTOR ANALYSIS CALCULATIONS BY TABULATING MACHINES - - - - - - - - - - - - - 93
D. M. HALL, E. L. WELKER, AND ISABELLE CRAWFORD

A NOTE ON RELIABILITY - - - - - - - - - - 127
HYMAN B. KAITZ

FACTORIAL DESIGN AND COVARIANCE IN THE STUDY OF INDIVIDUAL EDUCATIONAL DEVELOPMENT - - - - - - - - - - - - - - 133
PALMER O. JOHNSON AND FEI TSAO

ERRATUM - - - - - - - - - - - - - - - - 163

VOLUME TEN JUNE 1945 NUMBER TWO

PSYCHOMETRIKA—VOL. 10, NO. 2
JUNE, 1945

A MULTIPLE GROUP METHOD OF FACTORING THE CORRELATION MATRIX*

L. L. THURSTONE
THE UNIVERSITY OF CHICAGO

There are a number of methods of factoring the correlation matrix which require the calculation of a table of residual correlations after each factor has been extracted. This is perhaps the most laborious part of factoring. The method to be described here avoids the computation of residuals after each factor has been computed. Since the method turns on the selection of a set of constellations or clusters of test vectors, it will be called *a multiple group method of factoring*. The method can be used for extracting one factor at a time if that is desired but it will be considered here for the more interesting case in which a number of constellations are selected from the correlation matrix at the start. The result of this method of factoring is a factor matrix F which satisfies the fundamental relation $FF' = R$.

Selection of clusters

This method starts with the selection of a set of clusters of tests which are significantly correlated as in the original attempt to determine the dimensionality of a correlation matrix. It is assumed here that an estimate has been made for the diagonals. The grouping can be done systematically by listing all of the tests with high correlations in any given column of the correlation matrix. The rest of the square table for the group so selected is then completed from the correlation matrix. Those tests are eliminated from the group which have any low correlations so that the group which remains shows only significant correlations. If there are any negative signs, we reverse the sign of one or more variables so that the resulting correlations for the group are all positive. Then we have a cluster of tests whose vectors can be assumed to be roughly collinear in the sense that the test vectors of each group are contained in a cone with fairly small angular spread.

When one cluster has been selected, the next cluster may begin with a test which is not included in the first cluster. One proceeds in the same manner for each successive group. When all of the test vec-

* This study is one of a series of investigations in the development of multiple factor analysis and application to the study of primary mental abilities. We wish to acknowledge the financial assistance from the Social Science Research Committee of The University of Chicago which has made possible the work of the Psychometric Laboratory.

tors have been included in one or more clusters, it may be desirable to eliminate each test from all but one of the clusters. This should be the cluster in which the test has the highest average correlation. It is not desirable to do this grouping with any hair-splitting computations because clusters are rarely unique. For purposes of ultimate interpretation the clusters are frequently misleading because several clusters may be found in the test vectors of a single hyperplane. However, the dimensionality can be roughly estimated from the number of clusters that can be distinguished in the given correlation coefficients. The present purpose of arranging the tests in clusters is merely to define an oblique set of unit reference vectors which are linearly independent and which span the common factor space or as much of it as can be seen by inspection of the given correlations.

Computational procedure

In *Table 1* we have an illustrative example of nine variables and their correlations. Inspection of this table enables us to select three groups denoted s_1, s_2, and s_3 with several tests in each group. In this example there are three tests in each group, but in practice it is not necessary to have the same number of tests in each group. One starts the factoring by selecting the groups that seem to be conspicuous and there is no need to be concerned about whether the selection is in any sense unique. In general one should avoid assigning a test to two groups but that is not prohibited. Nor is it essential that all of the tests be represented in the subgroups although that was done in the present example.

In the matrix S we have three rows corresponding to the three groups. The tests in each group are listed in the table. In row s_1 we write the sum of the correlations for the tests in group s_1. Since there are three tests in this group, the row s_1 will have the sums of rows 2, 5, and 9 of the correlation matrix. The correlation matrix has now been augmented by an additional row, which does not alter its rank because this row is a linear combination of the given rows of the correlation matrix. If an additional column were added on the right side of the correlation matrix in the same manner, we should have an augmented correlation matrix of the same rank. This has not been done because it would be simply the transpose of the rows of matrix S which have already been computed. The row s_2 is computed in the same manner by summing rows 4, 6, and 7 of the correlation matrix. The row s_3 is computed by summing rows 1, 3, and 8 of the correlation matrix. These are the three groups into which the test battery has been divided by inspection of the correlation matrix.

The sum of the corresponding entries in each row of the matrix

S are then computed and recorded as shown in matrix T. The entry 4.930, for example, is the sum of the entries in row s_1 and in columns 2, 5, and 9. The other sums are determined in the same manner. This square matrix T of order 3×3 represents the lower right corner of the augmented correlation matrix in which the rank is unaltered.

TABLE 1

	1	2	3	4	5	6	7	8	9
1	.65	.00	.49	.28	.24	.66	.36	.63	.28
2	.00	.64	.24	.32	.48	.00	.00	.00	.56
3	.49	.24	.58	.12	.18	.42	.00	.63	.49
4	.28	.32	.12	.65	.66	.42	.63	.00	.28
5	.24	.48	.18	.66	.72	.36	.54	.00	.42
6	.66	.00	.42	.42	.36	.72	.54	.54	.24
7	.36	.00	.00	.63	.54	.54	.81	.00	.00
8	.63	.00	.63	.00	.00	.54	.00	.81	.36
9	.28	.56	.49	.28	.42	.24	.00	.36	.65

R_{jk}

	I	II	III
1	.234	.512	.577
2	.757	−.244	−.091
3	.410	.047	.640
4	.567	.530	−.215
5	.730	.376	−.215
6	.270	.678	.432
7	.243	.838	−.218
8	.162	.178	.867
9	.734	−.134	.306

F

	1	2	3	4	5	6	7	8	9
s_1	.52	1.68	.91	1.26	1.62	.60	.54	.36	1.63
s_2	1.30	.32	.54	1.70	1.56	1.68	1.98	.54	.52
s_3	1.77	.24	1.70	.40	.42	1.62	.36	2.07	1.13

S

	s_1	s_2	s_3
s_1	4.930	2.400	1.790
s_2	2.400	5.360	2.380
s_3	1.790	2.380	5.540

T

	W
w_1	.450377
w_2	.431934
w_3	.424859

	1	2	3	4	5	6	7	8	9
p_1	.234	.757	.410	.567	.730	.270	.243	.162	.734
p_2	.562	.138	.233	.734	.674	.726	.855	.233	.225
p_3	.752	.102	.722	.170	.178	.688	.153	.879	.480

V'

	s_1	s_2	s_3
p_1	2.220	1.081	.806
p_2	1.037	2.315	1.028
p_3	.760	1.011	2.354

U

s_1 : Tests 2,5,9
s_2 : Tests 4,6,7
s_3 : Tests 1,3,8

	p_1	p_2	p_3
p_1	1.000	.467	.342
p_2	.467	1.000	.437
p_3	.342	.437	1.000

$R_{pq} = F_{pm} F'_{pm}$

	I	II	III
p_1	1.000	0	0
p_2	.467	.884	0
p_3	.342	.314	.886

$F_{pm} = \Lambda'_{mp}$

	I	II	III
p_1	1.000	−.529	−.199
p_2	0	1.131	−.401
p_3	0	0	1.129

$(F'_{pm})^{-1} = \Lambda^{-1}{}_{mp}$

The next step is to compute the reciprocals of the square roots of the diagonal elements of the matrix T. These are the weights w_1, w_2, and w_3, which are shown to the right of the matrix T.

The matrix V' is computed by applying the weight w_1 to the first row of S, the weight w_2 to the second row of S, and so on. Similarly, the same weights on the rows of T give the square matrix U as shown. The matrix V' and U may be thought of as three rows by which the correlation matrix could be augmented instead of the rows of S and T. The matrix U would then be the lower right corner of the augmented correlation matrix in which the rank is still unaltered.

The three weights w_1, w_2, and w_3 are then applied to the *columns* of the matrix U and we then have the matrix R_{pq} with unit diagonals. The calculations are in part checked by the unit diagonals of this matrix. This matrix shows the cosines of the angular separations of the three unit reference vectors that have been chosen. Each one of them contains the centroid of a cluster of test vectors.

If we write the transpose of the matrix V' we have an oblique factor matrix V as in previous methods of rotation from the orthogonal factor matrix F, but we have obtained here the oblique matrix directly and without computing the orthogonal factor matrix F which is usually obtained first by the various factoring methods.

Since we have here an oblique factor matrix V as well as the cosines of the angular separations of the unit reference vectors for that factor matrix V, we can determine the transformation which should carry the oblique factor matrix V into an orthogonal factor matrix F. Since the location of an orthogonal frame in the system defined by V and U is not unique, we must first define in some convenient way the location of the desired orthogonal reference axes of F. We shall do this by locating the first of the orthogonal reference axes collinear with the first of the oblique axes. Then the first column of F will be the same as the first column of V. The next orthogonal reference axis will be located in the plane of the first two oblique axes and orthogonal to the first one. This process is in fact the diagonal method of factoring applied to the correlation matrix R_{pq}. When this is carried out routinely, we have the factor matrix $F_{pm} = \Lambda'_{mp}$ in which the three oblique reference axes p are defined in the orthogonal frame m.

The inverse of the transpose of this matrix gives $\Lambda_{mp}^{-1} = (F'_{pm})^{-1}$. We can now write the transformation by which the orthogonal factor matrix F can be written for the whole test battery. The matrix F_{pm} is then the additional three rows of F which describe the oblique reference axes p. This transformation is

$$V\Lambda^{-1} = F \tag{1}$$

from which we have

$$F\Lambda = V \qquad (2)$$

where Λ is already numerically given.

Applying this transformation to the matrix V by equation (1) we obtain the factor matrix F, which is shown in the upper right corner of the table.

When this factor matrix has been written, a residual table can be computed for the three factors to determine whether additional factors need to be extracted. In the present example, which is exact, the residuals vanish identically. In a practical problem with experimental data, the residual table should be computed and inspected to determine whether additional factors should be extracted. If that should be desirable, one proceeds with the residual table as with the original correlation matrix in the present example. Clusters are then listed from the residual correlation matrix until the tests with significant residuals seem to be represented. The process is repeated so as to add one or more columns to the factor matrix F.

In the present case we have one set of residuals to compute for the entire factor matrix F instead of three sets of residuals as would have been required in most of the current methods of factoring the correlation matrix. It will also be seen that the order of the inverse to be computed is equal to that of the number of factors determined by this process. For example, if seven factors are to be determined and if the process starts with seven linearly independent clusters, then there is an inverse of order 7×7 to be computed. If the factoring proceeds with, say, only four clusters, then there is an inverse of order 4×4 to be computed and the fourth factor residuals should then reveal three additional clusters resulting in three columns to be added to the factor matrix F. The computational labor is probably minimized by finding as many clusters in the given correlation matrix as there are common factor dimensions so that only one table of correlational residuals need be computed. However, if the first estimate of the number of clusters is too small, one merely repeats the process until the residuals vanish. If the factoring is done in several steps, the inverses are of course of smaller order.

The method of selecting groups for the present example was the method used in the early development of multiple factor analysis. It serves the purpose of separating the variables into a number of linearly independent groups that can be used for factoring as has been shown here. This method of group selection depends on the simple principle that those tests should be put in the same group which are highly correlated. If all of the test vectors were of equal length, i.e., if the communalities were all the same, then this principle would give a set of tests in each group that are roughly collinear. Since all of the

test combinations were not examined, it is to be expected that the groupings are not necessarily the best that could have been found.

The purpose of the grouping of the tests is to find sets that are roughly collinear. An improved method of accomplishing this result is to plot pairs of columns of the correlation matrix. Those tests which give roughly linear plots through the origin are nearly collinear and they should be placed in the same group. This method of grouping the tests is an improvement because it does not exclude a test of low communality from the group to which it should belong.

In using this method on the present nine-variable example, we start with the test whose column in the correlation matrix has the highest absolute sum. That is test 6. Plotting column 6 with the other columns we find that 1 and 6 are nearly collinear. These are placed in the first group. The next highest absolute sum is that of column 5. Plots with other columns show that 4 and 5 are nearly collinear. Eliminating these four tests, the next highest absolute sum is for test 9. Examining the plots for rough approximation to collinearity one could select the following groups: a) 1, 6, 8; b) 4, 5, 7; c) 2, 3, 9. As has been pointed out before, the grouping of tests is not unique. This grouping obtained by the graphical method could have been used in the present example and the result would have been the same because this example is exactly of rank 3. In dealing with experimental data it is best to select a number of groups that is larger than the probable number of common factors that can be interpreted. This insures that there will be at least one residual factor. This would not be successful on fictitious data of exact rank as in the present case because one would then find that the correlation matrix R_{pq} would not yield an inverse. Such a situation is rare with experimental data where it is preferable to have one or more residual common factors to be left without interpretation. The purpose of the residual factors is to insure that the common factor space has been covered with a sufficient number of dimensions of significant variance for interpretation.

In practice the grouping can be done by combining those tests which have roughly proportional columns of coefficients as judged by inspection. The graphs merely make the proportionality more obvious. Further, it is not necessary that all of the tests in each group be collinear or even nearly so. All that is required is that the whole correlation matrix be divided into sections that are linearly independent. This is accomplished by combining those tests which are significantly correlated.

PSYCHOMETRIKA—VOL. 10, NO. 2
JUNE, 1945

THE RELATION OF ITEM DIFFICULTY AND INTER-ITEM CORRELATION TO TEST VARIANCE AND RELIABILITY*

HAROLD GULLIKSEN
COLLEGE ENTRANCE EXAMINATION BOARD†

Under assumptions that will hold for the usual test situation, it is proved that test reliability and variance increase (a) as the average inter-item correlation increases, and (b) as the variance of the item difficulty distribution decreases. As the average item variance increases, the test variance will increase, but the test reliability will not be affected. (It is noted that as the average item variance increases, the average item difficulty approaches .50). In this development, no account is taken of the effect of chance success, or the possible effect on student attitude of different item difficulty distributions. In order to maximize the reliability and variance of a test, the items should have high intercorrelations, all items should be of the same difficulty level, and the level should be as near to 50% as possible.

The following proofs are restricted to cases where the test is composed of a number of items, each of which is counted as one point if correct and as zero if incorrect, and the raw score is the number of correct answers. The influence of chance success, and the possible effect on student attitude of various distributions of item difficulty are not considered. The effect of number of items on test parameters will not be pointed out, since this factor has been adequately dealt with by previous writers on test theory (see 4, 5, and 6).

Consider the relation of the item difficulty distribution to the maximum possible correlation between the item and total test score. In order to determine this relationship, we shall begin with the relation of item difficulty to the maximum correlation between two items. If N designates the number of persons taking the test, and p_i and p_j represent the proportions of persons answering items i and j correctly $(p_i \leqq p_j)$, then the sum of the cross-products of the scores on items i and j for the N individuals cannot exceed Np_j. Applying the usual gross score formula for correlation and simplifying we get

$$r_{ij}^{\max} = \sqrt{\frac{p_j(1-p_i)}{p_i(1-p_j)}}, \qquad (p_i \leqq p_j) \qquad (1)$$

* The desirability of determining this relationship has been indicated by previous writers. Work on the present paper arose out of some problems raised by Dr. Herbert S. Conrad in connection with an analysis of aptitude tests.

† On leave for Government war research from the Psychology Department, University of Chicago.

where r_{ij}^{max} is the maximum possible correlation between items i and j. Formula (1) or its equivalent is also given in reference (1), formula 19, and in reference (3), formula 3. Only when $p_i = p_j$ can r_{ij}^{max} equal unity. The maximum possible correlation declines very sharply as the difference between p_i and p_j increases. Figure 1 shows

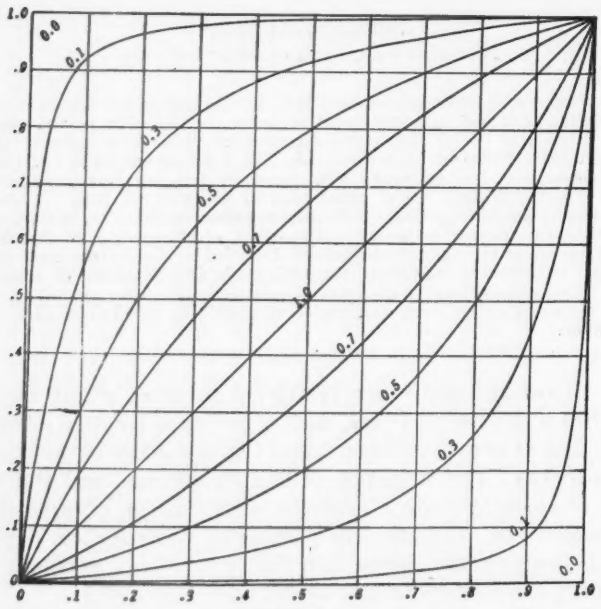


FIGURE 1

Combinations of p_h and p_k for which the Maximum Correlation is 1.0, 0.7, 0.5, 0.3, 0.1, and 0.0.

the difficulty combinations for which the maximum correlation is 1.00, .70, .50, .30, .10, and .00. In order to emphasize the sharpness of the drop in maximum inter-item correlation with increase in difficulty difference between two items, Figure 2 shows the highest possible correlation between two items for various values of p_h and p_k for the special case in which $p_h + p_k = 1$. As the difficulty difference between the two items increases, the drop in maximum correlation is very sharp. For example, an item answered correctly by 40% of the group cannot possibly correlate more than .667 with an item answered correctly by 60% of the group.

We can see from Figures 1 and 2 that only if all items are of the same difficulty value is it possible to have perfect correlation among

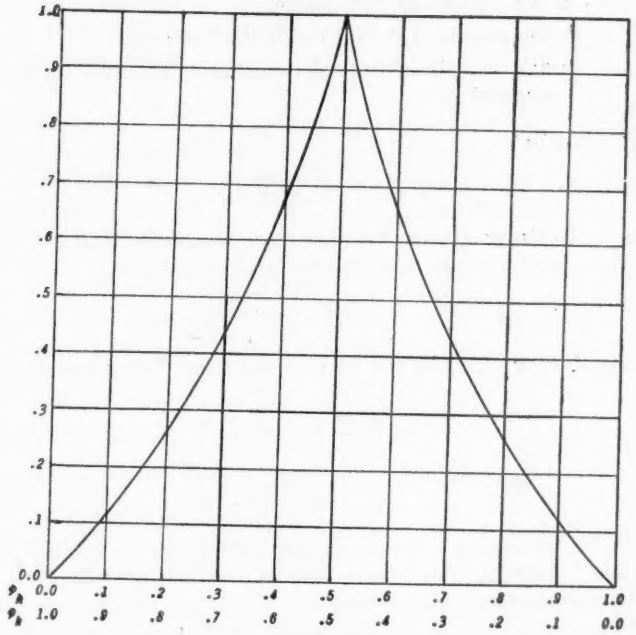


FIGURE 2
Maximum Inter-Item Correlation for Two Items of Difficulty p_h and p_k for the Special Case $p_h + p_k = 1.00$.

the items. Also, if the difficulty value is near 50%, then a difference in difficulty has less effect on $r_{ij}^{\max}$ than if the difficulty value is farther from 50%. It should be emphasized that this maximum is a *theoretical* one. It would be desirable to have some evidence for various types of items to show how great a difficulty difference it takes to become *practically* serious in decreasing inter-item correlation.

We may now see how the item difficulty distribution affects the maximum possible item-total correlation. Applying the formula for the correlation of a sum to a test of K items, we have

$$r_{it} = \frac{\sum_{j=1}^{K} r_{ij}\,\sigma_i\,\sigma_j}{\sigma_i\,\sigma_t}, \qquad (r_{ii} = 1) \tag{2}$$

where

r_{it} is the item-total correlation,

σ_t is the standard deviation of the test, and

σ_i and σ_j are the standard deviations of items i and j, respectively.

We also note that

$$\sigma_i \sigma_j = \sqrt{p_i q_i p_j q_j}, \tag{3}$$

where p_i is the proportion answering item i correctly, and $q_i = 1 - p_i$. Let us arrange the items in difficulty order such that

$$p_1 > p_2 > \cdots > p_i > p_j > \cdots > p_K .$$

Then substituting (1) and (3) in (2) and simplifying we obtain

$$r_{it}^{\max} = \frac{p_i \sum_{j=1}^{i} q_j + q_i \sum_{j=i+1}^{K} p_j}{\sigma_t \sqrt{p_i q_i}}, \tag{4}$$

where $r_{it}^{\max}$ is used to designate the maximum possible item-total correlation. Formula (4) above corresponds to formula 17 in reference (1) and formula 4 in reference (3). By arranging the items in difficulty order and running cumulative totals beginning at q_1 for the q's and at p_K for the p's, a routine for calculating (4) for each item of any given set of items can be set up. It can be seen that if all item difficulties (p_i and p_j) are equal, then all $r_{it}^{\max}$ are equal, and that the more the p_i differ from each other, the greater will be the spread in values of $r_{it}^{\max}$.

We may summarize the foregoing material in

Theorem A: The correlation between two items can be as high as unity only when they are of the same difficulty; as the difficulty difference between two items increases, the maximum possible correlation between them drops sharply.

Let us turn now to the effect of item difficulty on the mean, variance, and reliability of the total test score.

Since the mean of a sum equals the sum of the means of the separate parts, we may write

$$M_s = \sum_{i=1}^{K} p_i = K\bar{p}, \tag{5}$$

where

M_s = the mean raw score,
K = the number of items in the test,
p_i = the proportion of persons answering the i-th item correctly.
$\bar{p}$ = the average of the p_i;

[see also reference (1), formula 2, and reference (6), formula 22].

We may express the standard deviation of the raw score distribution by using the usual formula for the variance of a sum given in reference (5), page 213, formula 163. Applying this formula, we may write

$$\sigma_s^2 = \sum_{i=1}^{K} \sum_{j=1}^{K} r_{ij}\,\sigma_i\,\sigma_j, \qquad (r_{ii} = 1) \tag{6}$$

where

$$\sigma_i = \sqrt{p_i\,q_i}\,. \tag{7}$$

In these equations

σ_s^2 = the variance of the distribution of raw scores,
r_{ij} = the correlation between items i and j,
p_i = the proportion of persons answering the i-th item correctly,
K = the number of items in the test.

Formula (6) shows that the raw score variance will increase with each increase in correlation between any two items and will increase as the difficulty value of any item approaches .50.

For a *given average* item difficulty level, how do changes in the *variance* of the item difficulty distribution affect raw score variance? This relationship can be observed most easily if the form of formula (6) is changed. Let us use $r_{r(\sigma\sigma)}$ to designate the correlation between the intercorrelation of two items, and the product of their standard deviations. Some of the assumptions in the development which follows will be more reasonable if we omit from the correlation all terms which include $r_{ii} = 1$. This means summing the terms in the matrix of item variances and covariances, excluding the terms lying on the principal diagonal. We will designate this procedure by indicating the sum over $(K^2 - K)$ terms instead of over K^2 terms.

In order to indicate the sum over the $(K^2 - K)$ non-diagonal terms, let us adopt the definition

$$\sum_{i,j=1}^{K^2-K} r_{ij}\,\sigma_i\,\sigma_j = \sum_{i=1}^{K} \sum_{j=1}^{K} r_{ij}\,\sigma_i\,\sigma_j - \sum_{i=1}^{K} \sigma_i^2 \qquad (\text{where } r_{ii} = 1)\,.$$

Using this definition and applying the usual formula for correlation gives

$$r_{r(\sigma\sigma)} = \frac{\sum_{i,j=1}^{K^2-K} r_{ij}\,\sigma_i\,\sigma_j - (K^2-K)\overline{r_{ij}}\;\overline{\sigma_i\,\sigma_j}}{(K^2-K)\sigma_{r_{ij}}\sigma_{\sigma\sigma}}, \tag{8}$$

where

$\overline{r_{ij}}$ is the average inter-item correlation, (excluding $r_{ii}=1$)

$\overline{\sigma_i\,\sigma_j}$ is the average of the products of item standard deviations for all possible combinations such that $i \neq j$, and

the other symbols have the same definition as in equations (6) and (7).

From (8) and (6) we have

$$\begin{aligned}\sigma_e^2 &= \sum_{i,j=1}^{K^2-K} r_{ij}\,\sigma_i\,\sigma_j + \sum_{i=1}^{K}\sigma_i^2 \\ &= (K^2-K)\,[r_{r(\sigma\sigma)}\,\sigma_{r_{ij}}\sigma_{\sigma\sigma} + \overline{r_{ij}}\;\overline{\sigma_i\,\sigma_j}] + K\overline{(\sigma^2)}\,,\end{aligned} \tag{9}$$

where

$\overline{(\sigma^2)}$ is the average item variance, and the
other symbols have the same definitions as in previous equations.

To eliminate the term $\overline{\sigma_i\,\sigma_j}$ we adopt the definition

$$\sum_{i,j=1}^{K^2-K}\sigma_i\,\sigma_j = \sum_{i=1}^{K}\sum_{j=1}^{K}\sigma_i\,\sigma_j - \sum_{i=1}^{K}\sigma_i^2$$

and note that

$$\sum_{i=1}^{K}\sum_{j=1}^{K}\sigma_i\,\sigma_j = \left[\sum_{i=1}^{K}\sigma_i\right]^2 = K^2(\bar{\sigma})^2.$$

Substituting N times the mean for the "sum of scores" gives

$$(K^2-K)\overline{\sigma_i\,\sigma_j} + K\overline{(\sigma^2)} = K^2(\bar{\sigma})^2.$$

Then multiplying both sides by $\overline{r_{ij}}$, we have

$$\overline{r_{ij}}(K^2-K)\overline{\sigma_i\,\sigma_j} + \overline{r_{ij}}\,K\overline{(\sigma^2)} = \overline{r_{ij}}\,K^2(\bar{\sigma})^2, \tag{10}$$

where

$\overline{(\sigma^2)}$ is the average item variance, and is not to be confused with

$(\bar{\sigma})^2$ which is the square of the average item standard deviation.

We may substitute (10) in (9) and get

$$\sigma_s^2 = (K^2 - K) r_{r(\sigma\sigma)} \sigma_r \sigma_{\sigma\sigma} + \overline{r_{ij}} K^2 (\bar{\sigma})^2 + K \overline{(\sigma^2)} (1 - \overline{r_{ij}}), \quad (11)$$

where

K = the number of items in the test,

$r_{r(\sigma\sigma)}$ = the correlation of the correlation between two items with the product of their standard deviations (including only the non-diagonal elements in the correlation matrix)

σ_r = the standard deviation of the distribution of inter-item correlations (not including the terms $r_{ii} = 1$)

$\overline{r_{ij}}$ = the average inter-item correlation ($i \neq j$)

$\sigma_{\sigma\sigma}$ = the standard deviation of the products of the item standard deviations ($i \neq j$)

$(\bar{\sigma})^2$ = the square of the average item standard deviation

$\overline{(\sigma^2)}$ = the average item variance.

Let us consider the probable magnitude of these terms. As far as is known no direct evidence on the value of $r_{r(\sigma\sigma)}$ has ever been obtained, so that setting an upper bound for it is a guess. However, since there is an attempt to make r_{ij} as large as possible, it should have a small range in a well-constructed test, which would make for a low value of $r_{r(\sigma\sigma)}$. Let us assume that in a well-constructed test $r_{r(\sigma\sigma)}$ will not be much greater than 0.1. If we estimate that r_{ij} will seldom be higher than .3 and follow the policy of discarding items for which $r_{ij} < 0.1$, then 0.1 seems a plausible guess for the value of $\sigma_{r_{ij}}$. The maximum value of $\sigma\sigma$ is .25, and if there are very few items beyond the 10,90 difficulty level, the lower value of $\sigma\sigma$ is 0.09. With a range of 0.16, it seems that 0.1 is a reasonable upper bound for $\sigma_{\sigma\sigma}$. By making correspondingly reasonable approximations, we may say that $\overline{r_{ij}}$, $(\bar{\sigma})^2$, and $\overline{(\sigma^2)}$ will usually lie between 0.1 and 0.2. Since $r_{r_{(\sigma\sigma)}}$, $\sigma_{r_{ij}}$, and $\sigma_{(\sigma\sigma)}$ will each be small, their product will be still smaller, and hence the first term of (11) can usually be neglected. Neglecting the first term, we can see from equation (11) that the raw score variance of a test will increase with an increase in the number of items, the average inter-item correlation, the average item standard deviation, or the average item variance.

The effect of altering the dispersion of item difficulty while holding the average item difficulty constant will depend on the behavior of

$\overline{(\sigma^2)}$ and $(\bar{\sigma})^2$. First investigating $\overline{(\sigma^2)}$ we see that $\sigma_i^2 = p_i - p_i^2$ and hence

$$\overline{(\sigma^2)} = \bar{p} - \overline{(p^2)}. \tag{12}$$

We may denote the variance of the item difficulty distribution by σ_p^2 and from the usual formula for variance write

$$\sigma_p^2 = \overline{(p^2)} - (\bar{p})^2, \tag{13}$$

where

$\overline{(p^2)}$ is the average of the squared item difficulties, and $(\bar{p})^2$ is the square of the average item difficulty. If we substitute (13) in (12) we find that

$$\overline{(\sigma^2)} = \bar{p} - \overline{(p)^2} - \sigma_p^2, \tag{14}$$

from which it is clear that for a given average item difficulty the average item variance $\overline{(\sigma^2)}$ decreases as the variance of the item difficulty distribution increases. This means that as far as the last term of equation (11) is concerned, the test variance will increase as the variance of the item difficulty distribution decreases.

No simple method has been found for determining the variation of $(\bar{\sigma})^2$ with variation in σ_p for a given $\bar{p}$. It would seem reasonable, however, to find that $(\bar{\sigma})^2$ varied in the same general way as $\overline{(\sigma^2)}$. A rigorous proof of this fact can be given by writing

$$\begin{aligned} \sigma_i &= \sqrt{p_i(1-p_i)} = \sqrt{(.5+d_i)(.5-d_i)} \\ &= \sqrt{(.25-d^2{}_i)} = \sqrt{(.25-x_i)}. \end{aligned} \tag{15}$$

In equation (15) d_i is defined as $p_i - .5$, and $x_i = d_i^2$. The last expression in (15) may be expanded in a McLauren series, giving

$$\begin{aligned} \sigma_i = 0.5 &- x - \frac{2\cdot 1}{2!}x^2 - \frac{2^2\cdot 1\cdot 3}{3!}x^3 - \frac{2^3\cdot 1\cdot 3\cdot 5}{4!}x^4 \\ &- \frac{2^4\cdot 1\cdot 3\cdot 5\cdot 7}{5!}x^5 - \cdots - \frac{2^{n-1}\cdot 1\cdot 3\cdots(2n-5)(2n-3)}{n!}x^n. \end{aligned} \tag{16}$$

Applying Cauchy's ratio test for convergence we find that this series converges for $-1/4 < x < +1/4$, which means that the series converges for $0 < p < 1$. Summing terms like those in (16) and substituting d^2 for x, we have

$$\sum_{i=1}^{K} \sigma_i = K\bar{\sigma} = 0.5K - \sum_{i=1}^{K} d_i^2 - \sum_{i=1}^{K} d_i^4 - 2\sum_{i=1}^{K} d_i^6 - 5\sum_{i=1}^{K} d_i^8 - 14\sum_{i=1}^{K} d_i^{10} - \cdots, \tag{17}$$

the first three terms of which may be rewritten as

$$K\bar{\sigma} = 0.5K - K\hat{p}^2 - \sum_{i=1}^{K} e_i^2 - K\hat{p}^4 - 6\hat{p}^2\sum_{i=1}^{K} e_i^2 - 4\hat{p}\sum_{i=1}^{K} e_i^3 - \sum_{i=1}^{K} e_i^4, \tag{18}$$

where $\hat{p} = \bar{p} - 0.50$ and $e_i = p_i - \bar{p}$. Referring to equation (18) and the second term of equation (11), we see that for a given average item difficulty, the second term of equation (11) will increase as the dispersion of item difficulty decreases.* Thus we see that whenever the first term of equation (11) can be neglected, equations (14) and (18) show that for a given average item difficulty the test variance will increase as the variance of the item difficulty distribution decreases. In reference (4) it is stated, without proof, that test variance increases as item difficulty variance decreases.

We may summarize the foregoing results on test variance in

Theorem B:

For any test in which $\overline{r_{ij}} >> [\sigma_r\, r_{r(\sigma\sigma)}]$ *the raw score variance will increase as*

a. *the average inter-item correlation increases,*

b. *the average item standard deviation or variance increases (e.g., as the average item difficulty approaches .50)*

c. *the item difficulty variance decreases for any given average item difficulty*

It may be pointed out that the condition under which Theorem B is true will hold for any well-constructed test.

Let us now consider the effect of item difficulty on the reliability of the test. From formula (11), together with the Kuder-Richardson formula (2) given in reference (6), we may write

* It can be shown by a fairly lengthy argument (which is omitted here) that the effect of negative odd moments can be neglected since they are small in comparison with the even moments.

$$R = \frac{\sigma_s^2 - K\overline{(\sigma^2)} + \sum_{i=1}^{K} r_i \sigma_i^2}{\sigma_s^2}, \tag{19}$$

where

R = the reliability of the total test,
r_i = the reliability of item i, and

the other terms have the same meanings as in equation (11). If we designate the correlation between item reliability and item variance by $r_{r\sigma^2}$ and follow a procedure similar to that followed in equations (7) to (11) we obtain

$$R = \frac{\sigma_s^2 - K(\overline{\sigma^2}) + r_{r\sigma^2}\sigma_{r_i}\sigma_{\sigma^2}K + K\bar{r}_i(\overline{\sigma^2})}{\sigma_s^2} \tag{20}$$

where

$\bar{r}_i$ is the average item reliability,

$r_{r\sigma^2}$ is the correlation between item reliability and item variance,
σ_{r_i} is the standard deviation of the distribution of item reliabilities,
σ_{σ^2} is the standard deviation of the distribution of item variances, and the other terms have the same meanings as in equation (11). Again, assuming that we have a well-constructed test, it is reasonable to suppose that $r_{r(\sigma\sigma)}$, $r_{r\sigma^2}$, σ_{r_i}, and $\sigma_{r_{ij}}$ will each be small, hence their products will be still smaller, and terms involving these products may be neglected. If we neglect such terms, equation (20) may be rewritten and rearranged as

$$R = 1 - \frac{\overline{(\sigma^2)}(1 - \bar{r}_i)}{\overline{(\sigma^2)} + \bar{r}_{ij}[K(\bar{\sigma})^2 - \overline{(\sigma^2)}]}. \tag{21}$$

This equation shows that test reliability increases either with average item reliability or average inter-item correlation. To show the effect of item difficulty, (21) may be rearranged as

$$R = 1 - \frac{1}{\dfrac{(1 - \bar{r}_{ij})}{(1 - \bar{r}_i)} + \dfrac{K\bar{r}_{ij}}{(1 - \bar{r}_i)}\left[\dfrac{(\bar{\sigma})^2}{\overline{(\sigma^2)}}\right]}. \tag{22}$$

The effect of the term $\left[\dfrac{(\bar{\sigma})^2}{\overline{(\sigma^2)}}\right]$ can be seen by noting that $\sigma_\sigma^2 = \overline{(\sigma^2)} - (\bar{\sigma})^2$

Thus the ratio $\left[\frac{(\bar{\sigma})^2}{(\sigma^2)}\right]$ approaches a maximum value of 1.0 as the variation in item difficulty approaches zero. Hence it can be seen that, other things being equal, test reliability increases as the variance of the item standard deviations decreases. This result is also given in reference (2) for the special case in which all inter-item correlations are unity.

It is to be noted that the variance of the item standard deviations is lowered by concentrating items either around *one* difficulty level, or around any *two* difficulty levels p_i and p_j such that $p_i + p_j = 1.00$. For example, the variance of item standard deviation is zero if half the items are of 30% and half of 70% difficulty. However, as we have seen from Theorem A, inter-item correlations (and hence test reliability) can increase markedly as item difficulty differences decrease. Therefore, in order to increase test reliability, it is probably better to lower the variance of item standard deviations by concentrating all items about *one* difficulty level than by concentrating items about *two* difficulty levels whose sum is unity.*

It should be noted, however, that the effect (shown in formula 22) of the variance of item standard deviation upon test reliability is slight. For example, if half of the items had a difficulty value of .5 and the other half a difficulty value of .9, then the value of the ratio $\frac{(\bar{\sigma})^2}{(\sigma^2)}$ would equal 16/17, which differs but slightly from unity. It is to be noted, nevertheless, that if all items are concentrated about *one* difficulty level, then this slight effect reinforces the marked effect noted in Theorem A above; hence it seems clear that for a given *average* item difficulty, one is likely to get the best reliability by having all items of *that* particular difficulty.

The relationship shown in equations (21) and (22) may be stated qualitatively in

Theorem C: For a test in which $r_{r(\sigma\sigma)}$, $r_{r\sigma^2}$, σ_{r_i}, *and* $\sigma_{r_{ij}}$ *are negligible, the test reliability will increase with*

a. an increase in average item reliability,

b. an increase in average inter-item correlation,

c. a decrease in the variance of the item standard deviations.

* In reference (4) the increase in reliability with decrease in item difficulty variance is indicated as contingent upon the increase in test variance, and not upon the decrease in variance of item standard deviations.

It may again be pointed out that the condition under which Theorem C is true will hold for any well-constructed test.

As a guide to test construction, the material contained in Theorems A, B, and C may be summarized in the following statement. If we ignore the effect of chance success, and assume that we are dealing with a "well-constructed" test, then in order to maximize the reliability and variance of the test, items should have high intercorrelations, all items should be of the same difficulty level, and this level should be as near to 50% as possible.

One point may be noted where the results implicit in Theorems A and C [or in formulas (1) and (22)] are at variance with current test practice. These formulas indicate that if all the items are *concentrated* at one difficulty level, (even though that be a very high level of difficulty such as 20%, or a very low level of difficulty such as 80%) test reliability can be higher than is possible if the items cover a rather wide difficulty *range,* such as 10% to 50%, or 90% to 50%. Current practice in test construction favors spreading the items over a difficulty range rather than concentrating all items at one difficulty level. Current practice would be particularly opposed to concentrating all items at a very high or at a very low level of difficulty. Since the theory presented here does not allow for the influence of chance success, it does not allow for the fact that very difficult items would have a relatively much higher proportion of chance successes, and hence would necessarily have very low inter-item correlations, resulting in low test reliability. Also it may be that a test composed entirely of very difficult items would have an adverse effect upon student attitudes, and hence upon test reliability. However, neither of these considerations apply to a test composed entirely of very easy items. Whether it is actually best to concentrate all items at one difficulty level, in accordance with the theory presented here, or to distribute items over a difficulty range in accordance with present test practice, can be determined only by experiments such as those reported in (7), (8), and (11).

It may be, if items of graded difficulty levels are used, that counting one point for each item correct is not a proper scoring method. The score assigned should rather be a best estimate of the difficulty level reached, analogous to that used in the Binet test. Such a method of test scoring has been refined and applied in reference (10); however, the problem has not been generally solved for objective tests.

Another limitation in the theory here developed should be pointed out. The criterion of maximizing test variance cannot be pushed to extremes. Test variance is a maximum if half of the population makes zero scores, and the other half makes perfect scores. Such a

score distribution is not desirable for obvious reasons, yet current test theory provides no rationale for rejecting such a score distribution. Obviously the "best" test score distribution is one which accurately reflects the "true" ability distribution in the group, but there is perhaps little hope of obtaining such a distribution by the current procedure of assigning a score based upon sheer number of correct answers. At present the only solution to such difficulties seems to lie in some type of absolute scaling theory (9), to replace the "number correct" score.

REFERENCES

1. Carroll, John B. The effect of difficulty and chance success on correlations between items or between tests. *Psychometrika*, 1945, 10, 1-20.
2. Dressel, Paul L. Some remarks on the Kuder-Richardson Reliability coefficient. *Psychometrika*, 1940, 5, 305-310.
3. Ferguson, G. A. The factorial interpretation of test difficulty. *Psychometrika*, 1941, 6, 323-329.
4. Jackson, R. W. B. and Ferguson, G. A. Studies on the reliability of tests. Bull. No. 12 of the Dept. of Educ. Res., Univer. of Toronto, 371 Bloor St. West, Toronto 5.
5. Kelley, T. L. Statistical method. New York: Macmillan, 1924.
6. Kuder, G. F. and Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, 2, 151-160.
7. Richardson, M. W. The relation of difficulty to the differential validity of a test. *Psychometrika*, 1936, 1, 33-49.
8. Symonds, P. M. Factors influencing test reliability. *J. educ. Psychol.*, 1928, 19, 73-87.
9. Thurstone, L. L. A method of scaling psychological and educational tests. *J. educ. Psychol.*, 1925, 16, 433-451.
10. Thurstone, L. L. The scoring of individual performance. *J. educ. Psychol.*, 1926, 17, 446-457.
11. Thurstone, T. G. The difficulty of a test and its diagnostic value. *J. educ. Psychol.*, 1932, 23, 335-343.

PSYCHOMETRIKA—VOL. 10, NO. 2
JUNE, 1945

FACTOR ANALYSIS CALCULATIONS BY TABULATING MACHINES

D. M. HALL, E. L. WELKER, AND ISABELLE CRAWFORD*
UNIVERSITY OF ILLINOIS

I.B.M. tabulating equipment can be of considerable help in reducing the time and increasing the accuracy of multiple factor analysis, even if used for only a part of the calculations. Once the plugboard is wired and those cards punched which are used over and over again, problems involving any number of variables can be handled with dispatch. The correlation matrix is listed, the totals verified, and the signs changed on the tabulator. Then the factors and the residual coefficients are calculated by means of a calculator. Tucker's procedure has been modified by using a calculator instead of a multiplying punch, by reducing the number of cards used, by simplifying checks on calculations, by simplifying plugboard wiring, and by preparing work sheets on tabulator paper. Extraction of factors from 24 variables at the rate of one in four hour's time seems to justify the use of the tabulating equipment on small problems.

Multiple factor analysis is a valuable statistical technique, but the amount of work involved is almost prohibitive unless machine methods are employed. Tucker's procedure† demands the use of an 80-counter tabulator and a multiplying punch, two pieces of equipment available only in the larger tabulating installations. Our adaptation utilizes a 56-counter tabulator and a Monroe calculator Model AA-1, with negative multiplication and keyboard lock devices, and makes possible the calculation of factor problems involving over 100 variables with considerably less work than if the problem were performed on a calculating machine only.

This report consists of three parts, (I), a brief discussion of the fundamental factor theorems and equations, (II), the procedures for machine calculations of the factor loadings, and (III), an appendix summarizing instructions for each tabulation.

I. *The Factor Analysis Problem*

Suppose that 20 tests are given to 100 individuals. It is probable that the scores made by any individual are determined by his abilities, such as intelligence, manual dexterity, aptitude in numerical work,

* Assistant Professor of Agricultural Extension, College of Agriculture; Associate in Mathematics, College of Liberal Arts and Sciences; and Senior Office Appliance Operator, University of Illinois, Urbana, Illinois, respectively.

† Tucker, Ledyard R. The centroid method of factor analysis by punch cards. Microfilm Negative No. 1623, Department of Psychology, University of Chicago.

etc. It is conceivable that the score made by a given individual on any test could be predicted if the amount of his abilities pertinent to the test were known. This implies, of course, that the relating formula would be known also. Unfortunately, we seldom know all these facts. Consequently we must proceed on the theory that the correlation between the scores on the different tests is influenced by the abilities possessed by the individual tested.

If the same ability or abilities are required by each of two tests, then scores on these two tests will tend to vary together. Furthermore, as the number of common abilities increases and as the relative importance of these abilities in the two tests becomes the same, the more perfect will be the relationship between the scores made on the two tests.

Realizing these facts we can now state the problem of factor analysis as a simple question. Knowing the test scores only, can we reverse the process given above and discover the abilities involved and their importance in each of the tests? It may broaden our concept if we speak of the abilities or common causes as factors.

In a table of inter-correlation coefficients between test scores, it is often observed that certain groups of variables show relatively high inter-correlations with all variables within the group, and relatively lower ones between variables of the group and those outside. From the factor analysis viewpoint this situation would lead to the hypothesis that at least one factor is common to the tests within the group. Other groups unrelated to the first suggest the presence of other factors. Some members of a given group might belong to other groups through the medium of common factors. The factor analysis technique attempts to clarify this grouping phenomenon by measuring the tendency of each variable to join a cluster with other variables because of the presence of common factors.

1. *Assumptions and Notation.* Let K denote the number of variables to be analyzed, as K tests, for example. Let N individuals be scored $X_1, X_2, \cdots, X_K$ on these variables. Denote by r_{ij} the correlation between variables X_i and X_j. There will be $(K + 1)/2$ coefficients of correlation. Assume each variable is made up of certain factors as follows:

(1) an error factor due to imperfection in the instrument
(2) specific factors present in that variable only (for simplicity we assume only one)
(3) other factors present in varying amounts, each in at least one other variable in the group. For convenience of notation we shall assume only three of these are present. This will be seen to involve no restriction on the method.

To simplify formulas, let us assume that scores on all factors and all variables have been normalized so that all means are zero and all variances are unity. The following notation will be used:

Ei = score on the error factor of the ith variable,
Si = " " " specific " " " " " ,
x, y, z = " " " three factors of the ith variable, which occur in at least one other variable,
Z_i = score on the ith variable.

It is assumed that the factors are such that $Z_i = a_i x + b_i y + c_i z + d_i(Si) + e_i(Ei)$. The coefficients a_i, b_i, c_i, d_i, e_i are called factor loadings. They are relative weights applied to the additive factors, which are of course variable, and are used to determine a value of Z_i from known values of the five factors in any individual instance. This assumption of additivity is a restriction on the factors similar in nature to the assumption of linearity of regression in multiple correlation. The essential difference is that the identity of the variables on the right side is known in regression, but is to be discovered in factor analysis. In both cases the coefficients must be determined. We also assume that all factors are independent in the sense of being uncorrelated.

We shall now give a brief description of the Thurstone centroid method of factor analysis.

2. *Fundamental Formulas.* By normalizing the variables we made the following equations true:

$$\frac{1}{N}\sum_{i=1}^{i=N} x_i = 0; \qquad \frac{1}{N}\sum_{i=1}^{i=N} X_i^2 = 1; \qquad \frac{1}{N}\sum_{i=1}^{i=N} x_i y_i = 0.$$

Similarly, for all other factors and combinations of factors we made means to equal zero, variances to equal unity and correlations to equal zero.

The correlation, r_{12}, between Z_1 and Z_2 is given by

$$r_{12} = \frac{1}{N}\sum_{i=1}^{i=N} [a_1 x_i + b_1 y_i + c_1 z_i + d_1 (S1)_i + e_1 (E1)_i] \times [a_2 x_i + b_2 y_i + c_2 z_i + d_2 (S2)_i + e_2 (E2)_i].$$

If these expressions be multiplied we see that two types of terms are obtained; those with squares of factors, and those with cross-products. By hypotheses made above, all sums of squares divided by N are one, and all cross-products divided by N are zero. Hence the expression reduces to

$$r_{12} = a_1 a_2 + b_1 b_2 + c_1 c_2 .$$

Thus the correlation between two variables is given by the sum of the products of their loadings on common factors, since other correlations would result in the same form.

In a similar way the variance of Z_1, denoted by σ_1^2, is

$$1 = \sigma_1^2 = \frac{1}{N}\sum_{i=1}^{i=N} [a_1x_i + b_1y_i + c_1z_i + d_1(S1)_i + e_1(E1)_i]^2$$

$$= a_1^2 + b_1^2 + c_1^2 + d_1^2 + e_1^2 .$$

Let $h_1^2 = a_1^2 + b_1^2 + c_1^2 = 1 - d_1^2 - e_1^2$.

3. *Communality.* We call h_1^2, defined above, the communality of variable Z_1. It is that portion of the variance of variable Z_1 due to those factors common to Z_1 and one or more of the other variables, That is to say, we can think of it as that amount of the variance of Z_1 which can be explained in terms of variation in factors included in the other variables which are being studied. Stated negatively, communality is that portion of the variance not due to error and specific factors.

In practical work it is clear that the communalities will be unknown. They could be determined only if the factor loadings were known *a priori*. Hence they must be estimated. A number of methods might be given for this estimation, but only the simplest will be described. It will be noted in the derivation to follow that slight errors in such estimation will have little effect on the calculated loadings.

An upper bound to the communality of a test would be its reliability or self-correlation. Existence of specific factors would make the communality less than the reliability. On the other hand, we would usually expect the communality of a variable to be larger than any correlation coefficient using that variable with any other, sign being disregarded. This is due to the possibility of some specific factors. In rare cases it might be smaller. We use this approximate lower bound for estimating the communality. For variable Z_1 we pick the numerically largest correlation coefficient in which Z_1 is used. The communality of Z_1 is estimated as being this coefficient with positive sign.

4. *Finding the Factor Loadings of the First Factor.* Consider all correlation coefficients which involve Z_1, together with the communality of Z_1. Let r_{11} denote this communality.

$$\begin{aligned} r_{11} &= a_1a_1 + b_1b_1 + c_1c_1 \\ r_{12} &= a_1a_2 + b_1b_2 + c_1c_2 \\ r_{13} &= a_1a_3 + b_1b_3 + c_1c_3 \\ &\cdots\cdots\cdots \\ r_{1K} &= a_1a_K + b_1b_K + c_1c_K . \end{aligned}$$

Adding these, we obtain

$$\sum_{j=1}^{K} r_{1j} = a_1 \sum_{j=1}^{K} a_j + b_1 \sum_{j=1}^{K} b_j + c_1 \sum_{j=1}^{K} c_j.$$

Similarly,

$$\sum_{j=1}^{K} r_{2j} = a_2 \sum_{j=1}^{K} a_j + b_2 \sum_{j=1}^{K} b_j + c_2 \sum_{j=1}^{K} c_j$$

. . . . . .

and

$$\sum_{j=1}^{K} r_{Kj} = a_K \sum_{j=1}^{K} a_j + b_K \sum_{j=1}^{K} b_j + c_K \sum_{j=1}^{K} c_j.$$

If these be added, we obtain

$$\sum_{i=1}^{K} \sum_{j=1}^{K} r_{ij} = \left(\sum_{j=1}^{K} a_j \right)^2 + \left(\sum_{j=1}^{K} b_j \right)^2 + \left(\sum_{j=1}^{K} c_j \right)^2.$$

The double summation on the left includes each correlation twice and each communality once. The terms on the right are the squares of the sums of the factor loadings of the three factors common to two or more tests.

Assume the factor loadings to be determined in an order such that the sum of the loadings of the first factor is a maximum. In the notation above we wish to maximize $\sum_{j=1}^{K} a_j$. This will occur if $\sum_{j=1}^{K} b_j = \sum_{j=1}^{K} c_j = 0$. Some doubt exists as to the wisdom of this assumption. Other systems of factor analysis depend on different hypotheses. Since we are concerned only with the Thurstone system, no attempt will be made to criticize the basic theory.

The result of the above hypothesis is

$$\Sigma r = \sum_{i=1}^{K} \sum_{j=1}^{K} r_{ij} = \left(\sum_{j=1}^{K} a_j \right)^2,$$

where Σr denotes the double sum. From this we obtain $\sum_{j=1}^{K} a_j = \sqrt{\Sigma r}$. We assume the positive square root. This involves no loss of generality. It merely assigns a direction of meaning to factor one, as possession of an ability or lack of this ability.

Now

$$\sum_{j=1}^{K} r_{ij} = a_i \sum_{j=1}^{K} a_j$$

from equation (1), since

$$\sum_{j=1}^{K} b_j = \sum_{j=1}^{K} c_j = 0.$$

This gives

$$a = \frac{\sum_{j=1}^{K} r_{ij}}{\sum_{j=1}^{K} a_j} = \frac{\sum_{j=1}^{K} r_{ij}}{\sqrt{\sum r}}.$$

This is, then, a formula for the calculation of the loadings of the first factor. In other words, it says that the factor loading of a particular variable is a fraction with numerator given by the sum of all correlations involving that variable, including its communality, and denominator given by the square root of the sum of all correlations and communalities, each correlation entering twice, each communality once. Thus the numerator is the sum of a column of the correlation matrix, and the denominator is the square root of the sum of the column sums.

5. *Finding Residuals.* Return to the formula for the correlation coefficient $r_{ij} = a_i a_j + b_i b_j + c_i c_j$. We now know the values of the first factor loadings, the a's. Now we can solve for that portion of the correlation not dependent on factor one. This is called the first residual, denoted by $r_{ij \cdot 1}$ and given by the formula

$$r_{ij \cdot 1} = r_{ij} - a_i a_j .$$

The table showing these residuals is called the first residual table.

If we express $r_{ij \cdot 1}$ in the form $r_{ij \cdot 1} = b_i b_j + c_i c_j$, it is clear that the residuals are correlations between the variables after the effect of factor one has been removed. In this sense they parallel partial correlations.

6. *Necessity of Sign Change.* We now have in effect a new correlation matrix ready for analysis similar to that applied above. Direct repetition of the described procedure is not possible, however, since each column sum, and therefore the grand sum, of the residual table will be zero, except for rounding errors. Our hypothesis made above that $\sum_{j=1}^{K} b_j = \sum_{j=1}^{K} c_j = 0$ forces this result. Thus the formula which would parallel the one used above, namely,

$$b_i = \frac{\sum_{j=1}^{K} r_{ij\cdot 1}}{\sqrt{\sum_{i=1}^{K}\sum_{j=1}^{K} r_{ij\cdot 1}}},$$

is meaningless since both numerator and denominator are zero. This indeterminant form is evaluated by an artifice of sign changing.

7. *Sign-Change Procedure.* It is well known that changing the sign of a variable changes the sign of the correlation of that variable with another. It would also change the sign of factor loadings in the same way. We can change the column totals and grand total of the residual table by changing the signs of some of the variables. Consistent with our desire to have each factor removed to account for the largest possible portion of correlation, we wish to maximize column totals. This result is approximated in the following way. Delete diagonal entries, i.e., the communalities, from the residual table. The column sums now differ from zero. Change the signs of those variables which show negative column totals until the maximum positive total is obtained. While one cannot specify this process exactly, certain rules are helpful. Change the variables in a sequence of steps, not all at once. Do this in order of magnitude, changing the largest one (or the largest ones) first. Subsequent changes may indicate that a previous choice was unwise. In this case it should be reflected back. More specifically we would wish to reflect back to the original sign any variable which changed to a negative total on reflection of some other variable in a later step. This process will be discussed in more detail in section 3.

8. *Estimating Residual Communalities.* Although we have residual communalities calculated in the same way as residual correlations, it seems better not to use them, but to make new estimates at this time. Since original communalities were merely approximations, relative errors in residual communalities conceivably might be quite large. It is probable that the absolute value of the highest first residual in each column might be a more accurate figure. Hence this value is used. As was mentioned above, it is apparent that errors in communality estimates will have little effect on factor loadings.

9. *Finding Loadings on Factor Two.* We are now ready to apply to the residual matrix after sign change the same process for finding factor loadings on factor two as was used in finding factor one loadings from the original correlation matrix. From these loadings and the first residual table we proceed to other factors in the same

way. It is now clear that our assumption that our matrix contained three factors was not a restriction on the method.

10. *Changing Signs in the Original Correlation Matrix.* From the procedure explained above, we can see that it would be advantageous to apply the sign-change procedure to the original correlation matrix before removal of the first factor. This will enable the removal of the factor of maximum importance as defined by the method used.

11. *Rotation of Axes.* When we think of factors as dimensions in Euclidean space, each variable will determine a single point whose coordinates are the factor loadings on that variable. Thurstone uses a rotation of axes to obtain factors in which identification or naming is facilitated. Since we are proposing no modifications in this procedure, we shall not discus the problems involved in rotation. There are, however, many interesting aspects of the theory of rotation and its application to specific problems. The authors hope to consider this in later work.

II. *Procedure for Machine Calculation*

The procedure outlined below inclules the steps necessary to obtain the factor loadings for the centroid matrix.

1. *The Work Sheets.* In order to minimize errors and to facilitate calculations, work sheets were prepared on the tabulator. A portion of one is shown as Figure 1.* The original correlation matrix and each residual table were first laid out on these work sheets and the cards were punched from them. These sheets were laid out in units of 8×8 cells because a 56-counter tabulator will accommodate only 8 six-digit variables plus the card identification columns. Continuous tabulator paper will accommodate 3 such units per page.

Our problem included 79 variables and consequently 80×80 became the table limits. The work sheets carried variable numbers across the top printed from type 8 cards and variable numbers down the left side printed from type 0 cards. (See Codes in Table 1).

Because type 0, type 8, and type 7 cards may be used for any factor tabulation, sets of these cards to accommodate 100 variables were kept on file. We selected 80 type 0 cards and 80/8 sets of type 8 cards. The 10 type 8 cards for section 1 were sorted ahead of the 80 type 0 cards on (79-78)† and tabulated on 4-copy paper. Then the 10 type 8 cards for section 2 were substituted for those for section 1 and the pack was retabulated. This procedure continued until work

* All figures and tables are segregated at the end of this article.
† The number in parenthesis refer to card columns.

sheets for 10 sections were prepared. The wiring diagram is shown as Figure 2. Specific instructions for each operation are summarized in the Appendix.

One page of the work sheet upon which the correlation coefficients have been entered is shown as Table 2. The variable numbers across the top were printed from the type 8 cards, and those down the side from the type 0 cards.

1.1 *The correlation matrix.* The original correlation matrix was laid out by entering the correlation coefficients in their proper cells below the diagonal on one set of work sheets. The sum of the 79 coefficients for each variable was obtained. Then in order to avoid having to deal with minus coefficients, unity was added to each coefficient and a new table was laid out with each coefficient entered in its proper cell on both sides of the diagonal.

Thus a coefficient of .0334 became 01.0334 and one of —.0544 became 00.9456, its complement. All coefficients were punched from these work sheets as 6 digits. The zeros to the left of the decimals were not recorded merely to save work, as is shown in units 2 and 3 of Table 2.

It was found convenient in checking the sign-change procedure to carry a blank variable with all cell entries as 01.0000. We used variable 76 but would recommend that variable 1 be used in the future. If the number of variables is not a multiple of 8, the extra columns in the last section may be left blank but the last rows in all sections should be filled with values of 01.0000.

The communalities are not used at this time, and consequently the values in the diagonal cells from 1×1 to 80×80 are 01.0000. The estimated communalities will be added on the factor calculation sheets (operation 4). The type 1 cards are punched from these work sheets according to the code shown in Table 1.

2. *Coding and punching.* The color and corner cut of the eight different types of cards used are important to the mechanics of the operations. For the most part we retained the card number assigned by Tucker in order that reference may be made to his work. The table number (78-79) identifies all operations necessary to obtain a set of factor loadings. All cards and tabulation sheets from the original correlation matrix until the first factor loadings were calculated were identified as Table 1. Table 2 identified the residual table and the second sign-change operations, the calculation sheets, and all cards used to obtain the second factor loadings. The sample sheets shown in this paper are taken from Table 10 and consist of parts of the operations made in obtaining the tenth factor loadings.

2.1 *The type 0 cards.* These cards are a serially numbered set carrying variable number in (1-3) and unit number in (78-79). They were used in preparing the work sheets described in operation 1.

2.2 *The type 1 cards.* Type 1 cards were punched from the work sheets of $1 + r$ matrixes. The data in each row of each section were punched on separate cards in columns (11-58). Column 4 was sight-verified, then the cards were count-sorted on (3-1) to verify these columns.

2.21 *Tabulating the column totals.* In order to verify the calculation of the complements and the punching, the type 1 cards were tabulated. (Setup and operation of tabulator is described in the Appendix.) Column totals thus obtained were verified with the totals obtained in operation 1.1. These totals were used to calculate the column totals for the following residual table as described in operation 5.2 and to calculate one-half the sign change sum (operation 2.31).

2.22 *Listing the correlation table.* When the totals were verified and the cards corrected, the pack was counter-listed to reproduce the work sheets. A portion of the listing is shown as Table 3. (The right-hand column appended to this table will be described in 5.1).

2.3 *The type 6 cards.* The type 6 cards carry one-half the sign-change sums in fields identical to those used in type 1 cards. These sums are calculated from the column totals of Table 4.

2.31 *The sign-change sum.* To change the sign of a variable in a correlation matrix and adjust the column total accordingly, we need only subtract the variable twice from the column sum. Assume for the moment that Table 5 is a correlation matrix with communalities omitted and the diagonal cells filled with 01.0000, Table 5 is in fact a portion of our data with unity values subtracted by means of the type 7 cards.

Variable 18 shows the largest negative sum (—.1022) and is therefore selected to have its sign changed first. All coefficients in row 18 are subtracted from the totals twice and the remainders are the column sums with variable 18 changed to a positive sign.

But since the tabulator can be made to subtract only once during a card cycle, we may secure the same result by dividing the column total by 2 and subtracting the variable only once. This is made clear by the following equation;

$$\text{Sum} - 2r = 2\left(\frac{\text{sum}}{2} - \frac{2r}{2}\right).$$

Actually we need not double the difference on the right side of the equation because we need only relative values in order to select the proper variables to have their signs changed.

The column totals in Table 4 contain 80 unities which were added according to operation 1.1. Therefore we subtract 80 from each total before we divide by 2. No difficulty is experienced when the sums are positive, i.e., larger than 80.0000, but when they are negative, i.e, less than 80.0000 the calculation may be a bit confusing at first. Since we wish the difference to be in complement form, with 2 places carried to the left of the decimal, i.e., 100 larger than it actually is, and since we are to divide by 2, we merely add 200 to the column total before we subtract the 80. These operations are made clear by the following examples:

79.1130 variable 1, Table 4	80.2081 variable 6, Table 4
200.0000	
279.1130	
−80.0000	80.0000
2|199.1139	2| 00.2081
99.5565 = 1/2 SCE	00.1040 = 1/2 SCE

These 1/2 sign-change sums are shown on Table 4 just below the printed column totals and they were punched in 10 type 6 cards according to the code shown in Table 1. A new set of type 6 cards was punched for each residual table.

2.33 *The preliminary sign-change tabulation.* The type 6 cards were tabulated to permit us to select the variables with the largest negative sums for the first sign-change trial. In order that these negative sums will be in direct rather than complement form, the tabulator is wired to subtract cards carrying an x in column 7.* Since all type 6 cards carry x (7) the values they carry are reversed as they enter the counters and negative values are thus expressed in direct form.

Type 8 cards were sorted ahead of the type 6 cards on (6-5) and the preliminary sign-change tabulation shown as Table 6 was printed. Succeeding sign-change trials will be described as operation 3.

2.4 *The index cards.* Index cards were used to aid in selecting the cards for variables which were to be changed in signs. The variable number was printed by means of a numbering machine upon the tab and also punched in (1-3). Index cards cannot be punched except

* The x-control is sometimes known as the 11 punch.

by hand-feeding each card into a numeric key punch. Therefore, the numbers must be punched from the reverse side and an 001 was punched as 100. An index card for each variable was prepared.

2.5 *The type 7 cards.* Type 7 cards were used to subtract the one which was added to each r to make $1 + r$ values in the correlation tables. (See operation 1.1). These cards were used without alteration in each subsequent sign-change procedure.

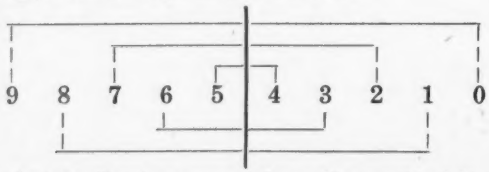
2.6 *The type 8 cards.* Type 8 cards were used in all tabulations to print the variable number as column headings.

2.7 *The type 9 cards.* Type 9 cards are punched to carry the communalities, i.e., the largest coefficient in each column. These were found by sorting the type 1 cards as follows:

2.71 Sort each detailed field (see codes) for each section on the first position to the left of the decimal, i.e., sort first on the column indicated under sort 1 below. Keep the 0's and 1's separate in subsequent sorting. Next, sort 1's on the column indicated under sort 2 and select cards in the highest sorter pocket. Sort 0's on the same column and select cards in the lowest sorter pocket.

Sort 1	Sort 2	Sort 3	Sort 4
Col. (12)	Col. (13)	Col. (14)	Col. (15)
Col. (18)	(19)	(20)	(21)
Col. (24)	(25)	(26)	(27)
Col. (30)	(31)	(32)	(33)
Col. (36)	(37)	(38)	(39)
Col. (42)	(43)	(44)	(45)
Col. (48)	(49)	(50)	(51)
Col. (54)	(55)	(56)	(57)

A plan for determining which of these pockets contains the highest valued r regardless of sign is given below:



The sorter pocket farthest removed from the vertical line should contain the cards carrying the highest r's. If "0" cards from the first sort are in pocket 2 and "1" cards are in pocket 6, only the cards in pocket 2 need be selected for subsequent sorting. If cards fall in pockets joined together in the diagram above both sets must be sorted on sort

3. If there are only 3 or 4 cards in these extreme pockets, they may be compared by converting the value on the "0" cards to complements and thus selecting the largest r. This value should be recorded on the table given below.

Estimated Communalities
Fields and Card Columns

Section	1(12-15)	2(18-21)	3(24-27)	4(30-33)
1				
2				
3				
etc.				

Section	5(36-39)	6(42-45)	7(48-51)	8(54-57)
1				
2				
3				
etc.				

The type 9 cards are punched from this table.

2.8 *The Type 10 Cards.* Type 10 cards were not punched until the sign-change procedure was completed. They contain a correction factor which is subtracted from the column totals after sign change to remove the unities which remain in the table. If the table contains 80 variables, 33 of which had their signs changed, 33×2 or 66 of the unity values would be cancelled out leaving $80 - 66 = 14$ unities to be taken out by the number 10 cards. This value will be found as the total for the check variable (76 in our case, see operation 1.1, and the tabulation sheet after sign-change, operation 3.1).

Cards "Index," 7, and 8 should be saved for future use; the other cards must be punched anew for each residual table.

3. *The sign-change procedure.* Prior to sign change, the index cards and the type 1 and type 7 cards were count-sorted in the order named on (3-1) and stored in a card file. The type 1 cards had already been verified (operation 2.21) and the type 7 cards were counted as they were sorted into the deck.

The preliminary sign-change trial (operation 2.33) merely permitted us to select the first variables to have their signs changed. These variables having been changed, a new tabulation was obtained

which permitted us to select the next group of variables to be changed in sign. This procedure was continued until all variables with column totals in direct form were changed in sign and all in complement form remained unchanged.

The variable with the largest column total was always selected first, and if there were several of about this same size, these were changed simultaneously. We have changed as many as 8 variables at one time during the early part of the procedure but toward the end it seems best to change only one at a time.

Following the preliminary sign-change trial variables 38, 48, 59, and 74 were circled on the tabulator sheet to indicate that they had been selected to have their signs changed because each of their column totals was considerably larger than any other. (See Table 6).

The cards for these variables were pulled from the card file and sorted on (6-5) behind the type 8 and type 6 cards taken from the tabulator after the tabulation of Table 6.

The index cards were rejected but the type 8, 6, and 7 cards were tabulated, using wiring diagram (see Figure 2). Minor control was (5-6) and (80); $x(7)$ was wired to subtract and suppressors remained as they were in operation 1.1.

The 3 steps in each sign-change trial are as follows:

1. One-half the column totals less the unities, carried on the type 6 cards enter the counters which were wired to subtract x (7).
2. To the counters were added the $1 + r$ values from type 1 cards.
3. From the counters were subtracted the unities by the type 7 cards carrying x in (7) and (10).

As a consequence the unities have been removed and the counters contain 1/2 the sums of r's from which has been subtracted the r's for the variables whose signs have been changed. All operations have been reversed in order that negative sums may be printed in direct form.

Using the next trial sheet after 38, 48, 59, and 74 have been changed, a second set of variables were selected, sorted, and tabulated as described above.

Table 7 was the 7th trial sheet upon which the variable numbers of those previously changed (shown in italics) were circled. It should be noted that variables 21 and 74 have been reversed. This may indicate that they should not have been changed, but because their numerical values are smaller than the values of 39, 51, and 61 we did not remove them at this time.

We used small circles about the variable number to show that the variable had been changed in a previous trial. We used a large circle about the variable number and its sum to indicate its sign was to be changed next; and we used a large X to indicate that the variable should be removed from the tabulated pack because in changing other variables the sign had been reversed. These symbols cannot be shown in type so bold face, italics, and underlining were used as indicated in the tables.

Table 8 shows variable 21 to be numerically larger and 74 smaller than on Table 7. Our problem was to decide whether to remove these variables and return them to the card file or to add new variables showing larger totals as do variables 5, 28, 50 and 78. We compared the true values, i.e., $5 = -.3741$, $21 = +.2208$, $28 = -.3145$, $50 = -.2428$, $74 = +.1233$, and $78 = -.4655$ and decided to remove 21, add 5, 28, and 78, and not remove 74 since its total is much smaller numerically than any of these. On trial 20 (Table 9) it is seen that this choice was justified.

Table 9 shows sign-change trial 20 for the tenth factor extracted in our problem. All variables in italics are larger than 0 and all not in italics are smaller. Since these values have entered the tabulator with $x(7)$ wired to subtract they are expressed as complements. Actually the signs of the variables whose column totals were negative have been changed to positive values, and the sign-change procedure is thereby completed.

3.1 *The sums after sign change.* The sums after sign change were now obtained directly from the type 1 cards in the following manner:

The cards which have been tabulated in operation 3 were sorted on (4) in order to separate them by type. This portion of the type 1 cards was gang-punched $x(9)$. These cards which were not tabulated were sorted on (4) and all index cards and types 6 and 7 were rejected.

The table totals needed in operation 4 should be obtained at this point, and before continuing with operation 3.1. All type 1 cards were tabulated with control on $x(9)$ subtract to obtain 8 field totals which were combined into a table total by hand, to equal item B_1t in Table 11a. The minor control switch was off. Then the type 9 cards were tabulated to obtain 8 field totals which were combined by hand into a table total to equal item Ct in Table 11a.

To continue with operation 3.1, the type 8 cards were sorted ahead of the type 1 cards on (6-5) and tabulated with control $x(9)$ subtract. Table 10 shows the column totals after sign-change and the table is checked by means of the check variable (76 in our case). All variables that have been changed must show larger totals than 76

shows and those not changed must be smaller. If more than half the variables had been changed the check column would be in complement form. The total of this check column is the amount that must be subtracted from each column total by means of the type 10 cards as explained in operation 2.8 and 4. Type 10 cards were then punched to carry the total for the check variable.

4. *The Factor Calculation Sheets.* The estimated communalities (see operation 2.7) were determined and the type 9 cards punched while the sign-change trials were being run.

As soon as operation 3 was completed, the following 4 cards were inserted after the last type 1 card in each section:

1. The type 10 card for the section.
2. A blank card.
3. The type 9 card for the section.
4. A blank card.

The pack was tabulated to provide the factor calculation sheet shown as Table 11.

The variables that have been changed were circled to indicate that all values for that variable and the factor loading will be negative. The rows in Table 11 are identified as follows:

B_1 row shows the sum of the columns for each variable after sign-change.

B_2 row shows the amount to be subtracted from each column total as is obtained from type 10 cards.

B_3 row is the difference between the B_1 and B_2 rows expressed in direct form. The values in this row must have the same sign as that given the variable. Negative numbers can be written in red.

C row contains the estimated communalities as obtained from the type 9 cards. The sign of this item is the same as that given the variable with negatives in red. This cell is left blank for any variable carried merely to fill out a section.

D row is the sum of the B_3 and C rows and each total must have the sign given the variable.

The calculations within Table 11 are checked in Table 11a. The B_1t total was obtained by summing the 8 field totals as described in operation 3.1. The B_2t total is the amount carried in any one field in the B_2 row times the number of variables in the table. The positive and negative values in the B_3 rows were summed separately. Their difference is the B_3t value in Table 11a, i.e., the table sum with regard

for signs, and it should equal the difference between the B_1t and B_2t totals in Table 11a. The positive and negative values in the C rows were summed separately, and their difference is the value of the Ct total. Likewise, the + and — values in the D rows are summed separately and their difference is the Dt total of Table 11a, and it should equal the sum of the B_3t total and the Ct total.

The square root of the sum of the r values for the table is obtained through Table 11b. The sum of the + and — values of the B_3 rows, i.e., the total without regard for signs is entered as B_3 total in Table 11b. The sum of the + and — values of the C rows equals the C total and the sum of the + and — values of the D rows equals the D total of Table 11b. $B_3 + C = D$ in this table as a check also. The square root of D is found and its reciprocal calculated. This value is used as a constant multiplier with each D value in Table 11 to obtain the factor loadings for each variable, which are entered to the left of the variable number in Table 11.

The sum of the factor loadings for the table should check except for rounding errors with the amount obtained by multiplying the sum of the D rows by the reciprocal obtained above.

5. *The residual tables.* A blank copy of the work sheets prepared in operation 1.1 was used for the calculation of the first residual table which was known as correlation Table 2. The sign of the loading of each variable on factor 1 was entered in the middle column; plus signs were entered in red pencil so that they could be clearly distinguished. (See Table 2.)

At this time a set of guide strips similar to the one shown to the right in Table 3 was prepared to carry the variable numbers and the loadings on the previous factor.

The residuals were then calculated by one worker using the Monroe calculator, while a second clerk recorded the values on the work sheets.

5.1 *Calculating the residual values.* The four guide strips were clipped to their respective pages of section 1 of Table 1 in such a manner that the variable numbers corresponded row by row with column 1.

The factor loading for variable 1 was locked in the Monroe as a constant multiplier. Since the signs of the factor loadings which are to be multiplied together must be carefully noted, we found it advisable to multiply all plus factor loadings in the column first, then return and multiply all minus loadings. The reason for marking the sign of the factor loading in each row on the work sheet is now apparent. The clerk recording must place the values in their proper cells.

By operating in this manner we needed to change the negative multiplication lever only once. The lever was set depending upon the sign of the factor loading locked in the machine. We wished to multiply 2 factors and subtract their product from the r of the previous table all in one operation, which is made more clear by the following examples.

r placed in lower dial	constant multiplier locked in		changing multiplier placed on keyboard		residual r taken from lower dial and copied on work sheet
.4217	− (−.10	×	−.10)	=	.4117
.3427	− (−.10	×	+.10)	=	.3527
99.3427	− (+.10	×	−.10)	=	99.3527
99.4217	− (+.10	×	+.10)	=	99.4117

The rules for operating the multiplication lever when one wishes to subtract the product of two factors from a coefficient were written on a card kept by the calculator.

1. Pull the lever forward if the signs of the terms being multiplied are different.
2. Push the lever backward if the signs of the terms being multiplied are alike.

To clear the key board with the 0 bar may change the setting of this lever, therefore, we always cleared by using the keys instead.

The recorder sat at the left of the operator and read the figures which appeared on the lower dials at the end of each multiplication. By the time the machine operator had set up the next coefficient upon the key board the recorder had written the former value upon the work sheets. The machine may thus be kept in continuous operation. Two operators have been able to calculate the 6400 residuals, punch the new set of cards and verify the calculations in 48 hours. The operators alternated positions frequently.

5.2 *Verifying the calculations and punching.* Generally it seems best to calculate the residual correlations column by column, and thus each is calculated twice and the table can be verified by reading columns against rows. As our operators became familiar with the procedure, we calculated only that half of the residual table below the diagonal. The values thus calculated were entered in the corresponding cells above the diagonal and sight-verified. The type 1 cards were punched from these work sheets after 01.0000 had been substituted for the value in each diagonal cell.These cards were sorted as described in operation 2.21 and tabulated to obtain section and unit totals as described in operation 2.22.

While the type 1 cards were being punched, the section totals for the next residual table were calculated in order that the cards might be verified as soon as the section totals were tabulated. The verification proceeded as follows:

The column total for variable 1 obtained from the original 1 plus r matrix was placed in the lower dial of the Monroe calculator. The sum of the loadings for factor 1 was locked in the Monroe as a constant multiplier. The factor loading for variable 1 was placed on the keyboard. The minus multiplication lever was set according to the rules cited in operation 5.1. The multiplication was then completed. However, since the communalities have been omitted from the type 1 cards and hence from the total, the square of the factor loading for variable 1 must be added. This is done by pulling down both the constant multiplier lever and the minus multiplication lever and squaring the factor loading on the keyboard. This figure must equal the column total of variable 1 as tabulated in operation 2.22 for the first residual table. To verify the total for column 2 the factor loading for variable 2 is used in the same way. This operation is similar to that of calculating the residual table. It differs in that column totals are used in place of individual coefficients and individual factor loadings, and in that an adjustment is made for having omitted the communalities in the type 1 cards. Each residual table is verified in this manner from the totals and factor loading of the preceding table.

If an error is found in a column total it can be isolated within a unit by applying this verification procedure to unit totals. Adjustment for communalities need not be made except in the unit containing the diagonal cell.

The factor loading tabulation sheet shown as Table 12 assists one in making the above calculations since it contains all the multipliers needed.

Appendix—Summary of Instructions for Each Tabulation

1.1 *The work sheets.*
Sort 10 type 8 cards for section 1 ahead of type 0 cards on (79-80). Tabulate using wiring diagram (see Figure 2). Add counter-list wires. Pull up short block-out levers 14-15, long hammer-lock levers for all type bars used and zero suppressers for 11, 16, 23, 29, 36, 43, 6, 13, 20, 27, 28. Set carriage to double-space and to print 9 4/6 inches on a 11-inch sheet, 4-copy paper. Sort out the type 8 cards on (80) and sort in the 10 type 8 cards for section 2. Tabulate as before. Repeat these operations for all 10 sections. Save cards for future use.

2.21 *Tabulating column totals.*
Use tabulator setting as in 1.1. Count-sort type 1 cards on (3-1).

Count-sort type 8 cards ahead of type 1 cards on (6-5). This is for verification. Tabulate with minor control on (5-6) and (80). One copy. Tabulate units total by changing control wires to (78-79) and (80).

2.22 *Listing the correlation matrix.*
Counter-list cards used in 2.21 with minor control on (78-79) and (80). The long hammer-lock levers are lowered for this operation. One copy.

2.33 *Preliminary sign-change tabulation.*
Sort type 6 cards behind type 8 cards on (6-5). Tabulate with minor control on (5-6) and (80) and x subtract control on (7). Pull up long hammer-lock levers again. One copy.

3.0 *Sign-change trial sheets.*
Select variables from tabulation 2.33 which show largest totals. Pull index cards, type 1 and type 7 cards from file box for these variables. Sort 8, 6, index, 1 and 7 cards in order given on (6-5). Tabulate with tabulator set as in 2.33.

3.1 *Sum after sign-change.*
Sort cards that have been tabulated in final trial of 3.0 on (4) in order to separate cards by type. Gang-punch $x(9)$ in this portion of the type 1 cards. Sort cards remaining in card file at end of operation 3.0 on (4). Combine the type 1 cards here with those that have been punched $x(9)$ and tabulate without minor controls (but change x subtract to (9)), in order to obtain 8 totals which may be summed into a table total. This is the B_1 total of Table 10b. Tabulate the type 9 cards without minor control in order to obtain C total of Table 10b.
Now sort type 8 cards ahead of type 1 cards on (6-5) and tabulate to sum after sign-change. This step may be omitted as it is the same as the B_1 row on the factor calculation sheets.

4.0 *Factor calculation sheets.*
We generally tabulate the type 1 and type 9 card totals as described above on the top of the calculation sheet.
The cards are arranged as follows for each section:

a. The type 8 cards
b. The type 1 cards
c. The type 10 card for the section
d. A blank card
e. The type 9 card for the section
f. A blank card

These are tabulated with minor control on (5-6) and (80) and $x(9)$ subtract.

FIGURE 1*

Work Sheets, for Entries of Original Correlation Matrix and Each Succeeding Residual Matrix

	1	2	3	4	5
1					
2					
3					
4					
5					
6					
7					
8					
	1	2	3	4	5
9					
10					
11					
12					
13					
14					
15					
16					
	1	2	3	4	5
17					
18					
19					
20					
21					

* Figure 1, Tables 3, 4 and 6-11 were set up on the tabulator.

TABLE 1

Codes for Cards Used

General layout	Card type							
	Index	0	1	6	7	8	9	10
Card form	Index	5080	5080	5080	5080	5080	5080	5080
Card color	Manila	Manila	Manila	Salmon	Green	Yellow	Salmon	Manila
Corner cut	Tab.	L.H.	L.H.	L.H.	L.H.	R.H.	L.H.	L.H.
No. of cards	80 1/	80	800	10	800	10	10	10
Col. No.								
(1-3) Variable No.	use	use	use	blank	use	blank	blank	blank
(4) Card No.	blank	0	1	6	7	8	9	blank
(5-6) Section No.	"	blank	use	use	use	use	use 2/	use
(7)	"	"	blank	x 3/	"J" 4/	blank	blank	blank
(8)	"	"	blank	x 5/	blank	"	"	"
(9)	"	"	x 6/	blank	"	"	"	"
(10)	"	"	blank	"	x 7/	"	"	"
			8/	9/				
Detail fields								
(11-16) Field 1	blank	blank	$1 + r$	use	blank	(15-16) 10/	use 11/	use 12/
(17-22) 2	"	"	$1 + r$	"	"	(21-22)	"	"
(23-28) 3	"	"	$1 + r$	"	"	(27-28)	"	"
(29-34) 4	"	"	$1 + r$	"	"	(33-34)	"	"
(35-40) 5	"	"	$1 + r$	"	"	(39-40)	"	"
(41-46) 6	"	"	$1 + r$	"	"	(45-46)	"	"
(47-52) 7	"	"	$1 + r$	"	"	(51-52)	"	"
(53-58) 8	"	"	$1 + r$	"	"	(57-58)	"	"
(76-77) Table No.	"	blank	use 13/	use	blank	blank	use	use
(78-79) Unit No.	"	use	use 14/	blank	"	use	blank	blank
(80)	"	omit	blank	"	"	1 15/	1	x (80) 16/

1/ This problem had 79 variables so 1 blank was entered to fill out the table.
2/ x (6) also to split control. 3/4/ x (7) subtract control. 5/ x (8) subtract.
6/ x (9) subtract only those variables that have had signs changed.
7/ x (10) subtract control for 1 in (7).
8/ Columns and decimal point identical with card 1.
9/ Use only 5th and 6th column of each field.
10/ Punch 0 in first column of each field.
11/ Decimal preceded by 00. 12/ Punch all 6 digits.
13/ Original table = No. 1, each residual table 1 unit higher.
14/ Number each 8 × 8 unit in each section. 15/ 1 for minor control.
16/ x (80) to split control.

TABLE 2

Work Sheets Showing Residual Correlations for a Part of the Table for Factor 10. (The original entries in this table were made by hand on a work sheet but are set in type here. See operation 1.1)

	1	2	3	4	5	6	7	8
49	010334	009456	009279	010604−	009856	009908	011154	010682
50	010071	009810	008254	009608+	010568	010313	009192	009675
51	010739	010290	009995	009979—	010008	010080	008777	007804
52	010040	010772	010529	011325—	010804	010604	010140	010011
53	009010	009086	009761	009572−	010134	009368	009175	008347
54	009325	010770	008196	009150−	010131	011042	010911	010417
55	010353	010048	010169	011300+	010430	010273	010176	010927
56	008933	010770	011148	013047+	010658	009940	010384	010673
	1	2	3	4	5	6	7	8
57	8515	9232	9122	8806−	10271	9760	9409	10562
58	9410	11518	9098	8851+	9576	10516	9453	9783
59	9937	11421	10719	5983+	9234	10584	10960	10855
60	9215	8936	10119	9941+	10891	10324	10252	10621
61	9887	9137	9832	5854+	11832	9791	11281	11967
62	9878	10354	8137	10914+	9634	9309	9677	10712
63	11258	10770	9763	11656−	9324	10173	8923	8470
64	8120	7261	10038	9751−	9064	10351	10131	10984
	1	2	3	4	5	6	7	8
65	10169	8999	8270	10602+	11245	10290	8794	9580
66	9863	10696	11128	9445+	9007	9595	11192	8523
67	8145	9852	10124	10601—	10093	10884	10965	11497
68	9451	9667	9458	11249−	10721	10309	10910	10629
69	11095	8567	11010	10836+	11228	10073	10412	9594
70	10935	7810	10124	12080+	9890	9569	10024	10347
71	10203	9666	10267	11877−	9588	10228	9599	9172
72	9945	9482	9430	11438+	9187	10350	10251	10475

TABLE 3

The Tabulation of Work Sheets Shown in Table 2 with Unit Totals as the 8th, 16th, and 24th Rows. A strip clipped to this table was used in the calculation of the residual value. (This table was listed on the tabulator but is shown here set in type. The figures to the right were entered by hand on a strip of tabulator paper so that the entries matched exactly row by row with the tabulated sheet.)

								Variable No.	Leading Factor 9
49	01	10334	9456	9279	10604	9856	9908	49	—.03
		10071	9810	8254	9608	10568	10313	50	+.13
		10739	10290	9995	9979	10008	10080	51	—.15
		10040	10772	10529	11325	10804	10604	52	—.32
		9010	9086	9761	9572	10134	9368	53	—.16
		9325	10770	8196	9150	10131	11042	54	—.05
		10353	10048	10169	11300	10430	10273	55	+.25
		8933	10770	11148	13047	10658	9940	56	+.24
49	01	78805	81002	77331	84585	82589	81528		
57	01	8515	9232	9122	8806	10271	9760	57	—.06
		9410	11518	9098	8851	9576	10516	58	+.10
		9937	11421	10719	5983	9234	10584	59	+.20
		9215	8936	10119	9941	10891	10824	60	+.15
		9887	9137	9832	5854	11332	9791	61	+.21
		9878	10354	8137	10914	9634	9309	62	+.08
		11258	10770	9763	11656	9324	10173	63	—.18
		8120	7261	10038	9751	9064	10351	64	+.12
57	01	76220	78629	76828	71756	79826	80808		
65	01	10169	8999	8270	10602	11245	10290	65	+.08
		9863	10696	11128	9445	9007	9595	66	+.26
		8145	9852	10124	10601	10093	10884	67	—.34
		9451	9667	9458	11249	10721	10309	68	—.35
		11095	8567	11010	10836	11228	10073	69	+.17
		10935	7810	10124	12080	9890	9569	70	+.22
		10203	9666	10267	11887	9588	10228	71	—.14
		9945	9482	9430	11438	9187	10350	72	+.21
65	07	79806	74739	79811	83138	80959	81298		

TABLE 4

The Column Totals for Each Variable Tabulated from the Cards. [Column totals shown in ordinary type were printed on the tabulator. Those in bold face (top row) were obtained in operation 5.2 and were entered by hand. Those in italics (lower row) were obtained in operation 2.31 and entered on this sheet by hand in red.]

	1	1	2	3	4	5	6	7	8
		791130	**797143**	**788973**	**797278**	**796099**	**802081**	**806854**	**798551**
3	01	791130	797143	788973	797278	796099	802081	806854	798551
		995565	*998571*	*994486*	*998634*	*998049*	*001040*	*003427*	*899275*
	2	9	10	11	12	13	14	15	16
		798256	**810913**	**799208**	**793541**	**805830**	**796145**	**789734**	**799457**
3	02	978256	810913	799208	793541	805830	796145	789734	799457
		999128	*005456*	*999604*	*996770*	*002915*	*998072*	*994867*	*999728*
	3	17	18	19	20	21	22	23	24
		792270	**801795**	**793850**	**812386**	**788522**	**790187**	**799368**	**796881**
3	03	792270	801795	793850	812386	788522	790187	799368	796881
		996135	*000897*	*996925*	*006198*	*994261*	*995093*	*999684*	*998440*
	4	25	26	27	28	29	30	31	32
		799882	**798717**	**799936**	**801287**	**801781**	**798286**	**809906**	**805479**
3	04	799882	798717	799936	801287	801781	798286	809906	805479
		999941	*999358*	*999968*	*000643*	*000890*	*999143*	*004953*	*002739*
	5	33	34	35	30	37	38	39	40
		800409	**801460**	**807703**	**794729**	**805190**	**780331**	**790696**	**795173**
3	05	800409	801460	807703	794729	805190	780331	790696	795173
		000204	*000730*	*003851*	*997364*	*002595*	*990165*	*995348*	*997586*
	6	41	42	43	44	45	46	47	48
		814152	**800176**	**814646**	**808492**	**793809**	**806959**	**818154**	**786270**
3	06	814152	800176	814646	808492	793809	806959	818154	786270
		007076	*000088*	*007323*	*004246*	*996904*	*003479*	*009077*	*993135*
	7	49	50	51	52	53	54	54	56
		794104	**801744**	**790932**	**796805**	**795528**	**806263**	**810106**	**802094**
3	07	794104	801744	790932	796805	795528	806263	810106	802094
		997052	*000872*	*995466*	*998402*	*997764*	*003131*	*005053*	*001047*
	8	57	58	59	60	61	62	63	64
		803008	**790526**	**782428**	**792551**	**780322**	**791895**	**800029**	**799257**
3	08	803008	790526	782428	792551	780322	791895	800029	799257
		001504	*995263*	*991214*	*996275*	*990161*	*995947*	*000014*	*999628*
	9	65	66	67	68	69	70	71	72
		799460	**808599**	**811634**	**816528**	**792378**	**812788**	**794609**	**788136**
3	09	799460	808599	811634	816528	792378	812788	794609	788136
		999730	*004299*	*005817*	*008264*	*996189*	*006394*	*997304*	*994068*
	10	73	74	75	76	77	78	79	80
		809367	**785111**	**791545**		**788318**	**795486**	**789841**	**803859**
3	10	809367	785111	791545	800000	788318	795486	789841	803859
		004683	*992555*	*995772*		*994159*	*997743*	*994920*	*001929*

TABLE 5

A Portion of Table 3 with the Unity Values Subtracted

Variables	15	16	17	18	19	20	21
15	.0000	.0934	—.1203	.1719	.0358	.0120	—.0099
16	.0934	.0000	—.0174	—.1951	.1322	.0646	.0057
17	—.1203	—.0174	.0000	.0460	—.0019	.0463	—.0481
18	.1719	—.1951	.0460	.0000	—.0982	—.0825	.0557
19	.0358	.1322	—.0019	—.0982	.0000	—.0525	—.0486
20	.0120	.0646	.0463	—.0825	—.0525	.0000	.1595
21	—.0089	.0057	—.0481	.0557	—.0486	.1595	.0000
Sum	.1829	.0834	—.0954	—.1022	—.0332	.1474	.1143
Calculations when 2 r_{18} is subtracted from the column totals							
Sum	.1829	.0834	—.0954	—.1022	—.0330	.1474	.1143
2 r_{18}	.3438	—.3902	.0920	.0000	—.1964	—.1650	.1114
Diff.	—.1609	—.3068	—.0034	—.1022	+99.8368	—.0176	+.0029
Calculations when r_{18} is subtracted from 1/2 the column totals							
1/2 Sum	.0914	.0147	—.0477	.0511	—.0166	.0727	.0571
r_{18}	.1719	—.1951	.0460	—.1022	—.0982	—.0825	.0557
Diff.	—.0805	—.1534	—.0017	—.0511	+9.9174	—.0088	.0014
× 2 for check	—.1610	—.3068	—.0034	—.1022	+99.8348	—.0176	+.0028

TABLE 6

Preliminary Sign-Change Trial Using Only Type 8 and Type 6 Cards. (See operation 2.33. These entries were tabulated and the figures in bold face were circled to show the variable was changed in sign.)

1	1	2	3	4	5	6	7	8
1	4435	1429	5514	1366	1951	998960	996573	725
2	9	10	11	12	13	14	15	16
2	872	994544	396	3230	997085	1928	5133	272
3	17	18	19	20	21	22	23	24
3	3865	999103	3075	993807	5739	4907	316	1560
4	25	26	27	28	29	30	31	32
4	59	642	32	999357	999110	857	995047	997261
5	33	34	35	36	37	38	39	40
5	999796	999270	996149	2636	997405	9835	4652	2414
6	41	42	43	44	45	46	47	48
6	992924	999912	992677	995754	3096	996521	990923	6865
7	49	50	51	52	53	54	55	56
7	2948	999128	4534	1598	2236	996869	994947	998953
8	57	58	59	60	61	62	63	64
8	998496	4737	8786	3725	9839	4053	999986	372
9	65	66	67	68	69	70	71	72
9	270	995701	994183	991736	3811	993606	2696	5932
10	73	74	75	76	77	78	79	80
10	995317	7445	4228		5841	2257	5080	998071

ALPHABETIC ACCOUNTING MACHINE PLUGBOARD

TYPE 405 WITH NET BALANCE COUNTERS AND AUTOMATIC CARRIAGE PLUGGABLE CONTROLS

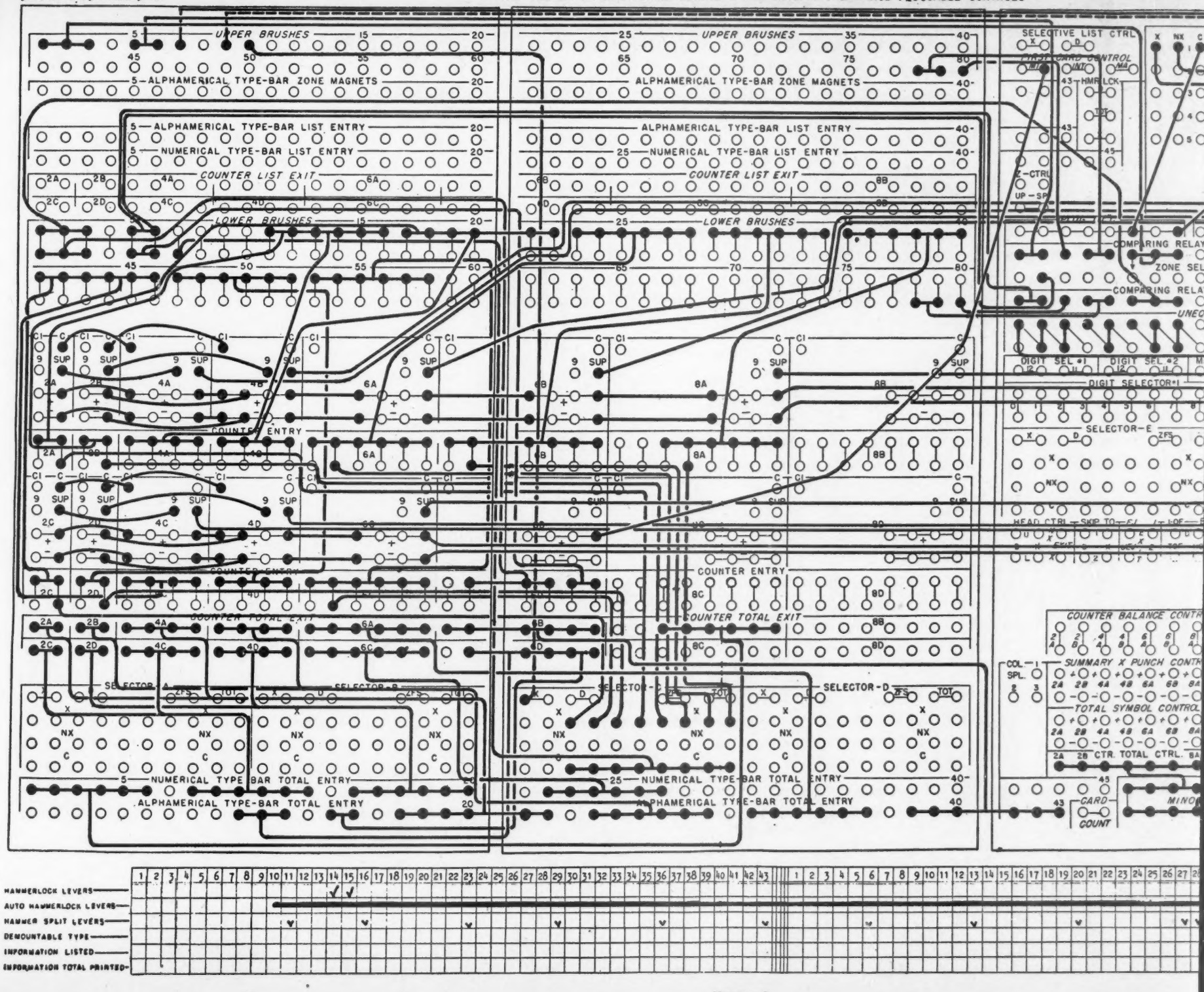


FIGURE 2

Wiring Diagram for Tabulating Factor Analysis Problems. (Counter list wires are not shown. Switches are set as follows: Automatic control mino Final total - - on, Automatic total minor - - on, Automatic start - - on, Hammerlock - - on, Selective list - - on, and the three list-tabulate switches accordi operation being performed.)

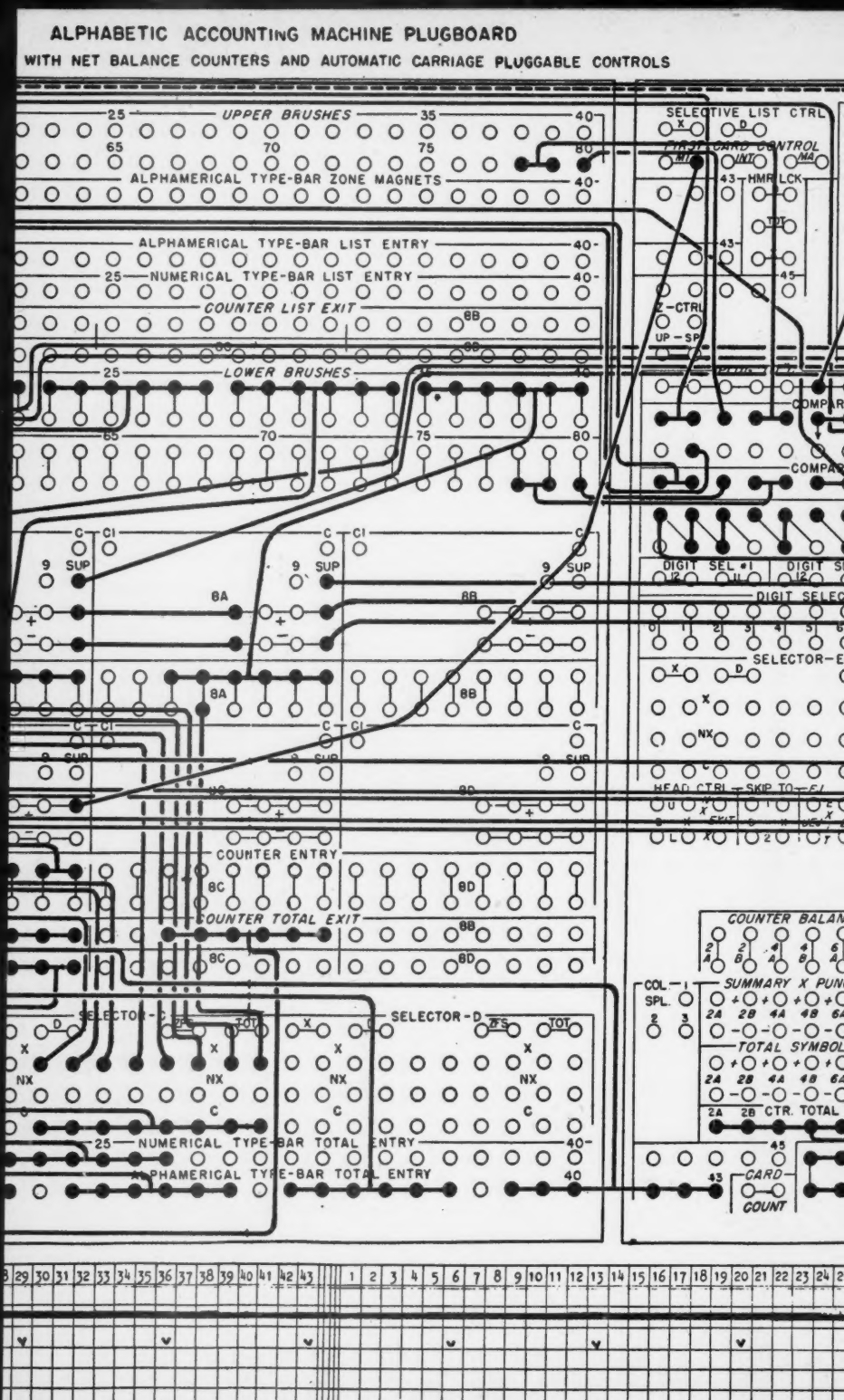
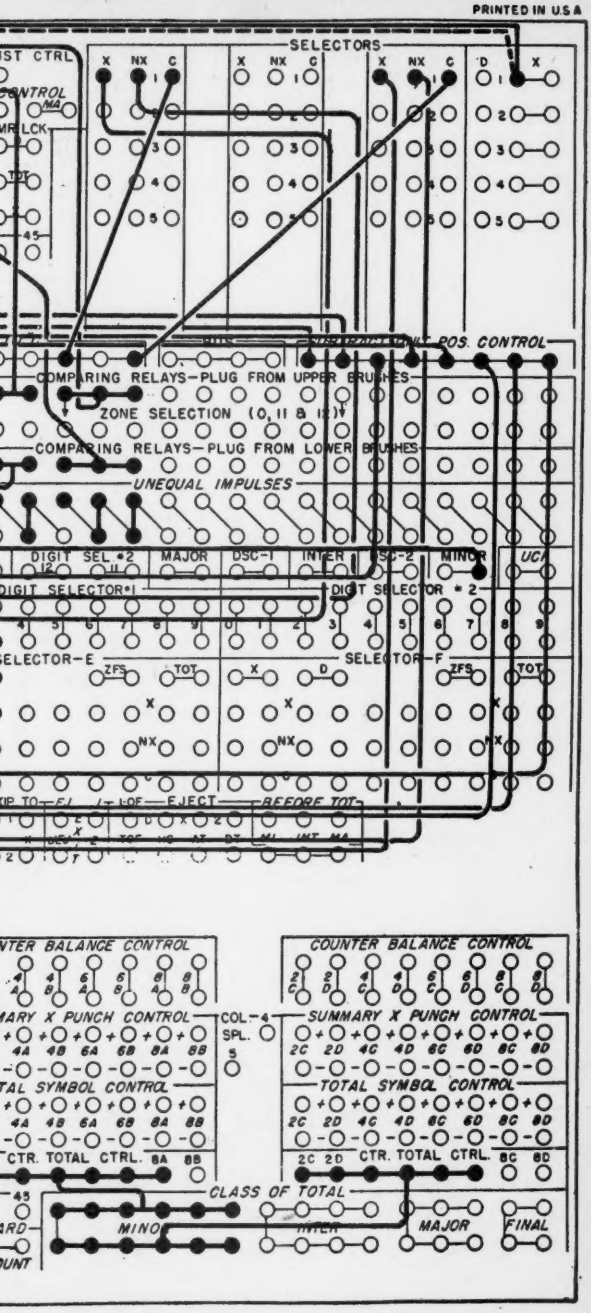
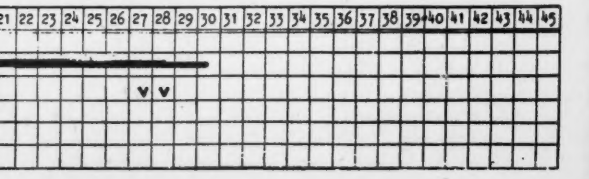


FIGURE 2

ysis Problems. (Counter list wires are not shown. Switches are set as follows: Automatic contr
omatic start - - on, Hammerlock - - on, Selective list - - on, and the three list-tabulate switches a

PRINTED IN U.S.A.





matic control minor - - on,
switches according to the

TABLE 7

Trial Sheet upon Which Variable Numbers of Those Previously Changed Have Been Circled. Variables 39, 51, and 61 are to be Changed Next. (These entries were all tabulated, and the variable numbers in italics were circled to show that they had been changed, while those in bold face were circled to show that they were to be changed in the next operation.)

1	1	2	*3*	4	5	6	*7*	*8*
1	997757	997375	9751	988267	3741	997306	5426	4385
2	9	10	11	*12*	13	14	*15*	16
2	994942	990982	993407	4491	995792	994180	7015	995071
3	17	18	19	20	*21*	22	23	24
3	993371	1334	999166	996807	997792	2137	995149	55
4	25	26	27	28	29	*30*	31	32
4	879	1534	993698	3145	991230	6716	991640	994745
5	33	34	35	*36*	37	*38*	39	40
5	999686	996661	991319	4391	997525	10540	3949	985676
6	41	42	43	44	45	46	47	*48*
6	988286	992094	994226	997414	993129	991348	987414	4976
7	49	50	51	52	53	54	55	56
7	993760	2428	3932	995396	996753	994633	996254	369
8	57	58	*59*	*60*	61	62	63	64
8	997770	992685	5356	5568	4074	996332	994793	1367
9	65	66	67	68	*69*	70	71	72
9	993983	994867	995775	990254	5641	996794	998568	1883
10	73	*74*	*75*	76	*77*	78	*79*	80
10	999244	998767	1975		8552	2824	7157	992599

TABLE 8

The 8th Trial Sheet; Variables 5, 28, and 78, Are to Have Their Signs Changed Next (These entries were tabulated originally. Then the variable numbers in italics were circled to show that they had been changed previously. Those in bold face, that is, 5, 28, and 78, were enclosed in a large circle to show that they were to be changed next, and the figures underlined were crossed out with a large X to indicate that the cards for 21 were removed because changes in other variables had reversed its sign.

1	1	2	*3*	4	5	6	*7*	*8*
1	995197	994346	8109	987039	5293	997270	5323	4502
2	9	10	11	*12*	13	14	*15*	16
2	992861	989001	992523	6003	997849	994949	9122	994894
3	17	18	19	20	*21*	22	23	24
3	994264	2342	341	992514	994227	1382	995922	613
4	25	26	27	28	29	*30*	31	32
4	1892	276	993028	4675	991381	5644	992665	994828
5	33	34	35	*36*	37	*38*	*39*	40
5	999422	995162	991612	5699	2485	10846	4135	984991
6	41	42	43	44	45	46	47	*48*
6	984999	994390	994338	997069	992649	992936	987613	7242
7	49	50	*51*	52	53	54	55	56
7	991891	2467	5532	993631	998251	994825	999521	998437
8	57	58	*59*	*60*	*61*	62	63	64
8	999116	991933	5353	10381	5810	996754	994371	406
9	65	66	67	68	*69*	70	71	72
9	998200	993838	994598	992210	7795	995334	75	997215
10	73	*74*	*75*	76	*77*	78	*79*	80
10	996241	999233	2814		8095	4655	8681	989097

TABLE 9

The 20th Sign-Change Trial. All Variables Shown in Italics and Whose Signs Should Have Been Changed Are Now Changed.

1	1	2	*3*	4	*5*	6	*7*	*8*
1	991932	995006	651	985047	3891	994651	6219	4277
2	9	10	11	*12*	*13*	14	*15*	16
2	988487	996601	986603	11887	406	993076	11130	998667
3	17	*18*	*19*	20	21	*22*	23	*24*
3	994914	1900	4147	993881	989590	3501	998010	6207
4	*25*	26	27	*28*	29	*30*	31	32
4	9798	992322	995860	9948	993306	5972	994394	994380
5	33	34	35	*36*	*37*	*38*	*39*	40
5	994327	994842	992252	2274	6680	10477	4175	988712
6	41	42	43	*44*	45	46	47	*48*
6	986153	996977	998431	3323	988256	997944	989077	2522
7	49	*50*	*51*	52	53	*54*	55	56
7	986623	6979	8124	992999	992774	1546	996699	992751
8	57	58	59	*60*	*61*	62	63	*64*
8	998448	991792	997578	11049	6870	992106	999695	3702
9	*65*	66	67	68	*69*	70	71	72
9	6667	990699	992490	992685	4077	994560	997643	997219
10	73	*74*	*75*	76	*77*	*78*	*79*	80
10	995097	3935	14351		2547	9308	10397	983533

TABLE 10

The Column Sums after the Sign-Change Procedure Has Been Completed. (The variables shown in italics have had their signs changed.)

	1	1	2	*3*	4	*5*	6	*7*	*8*
1	01	156136	149989	138699	169916	132219	150699	127562	131447
	2	9	10	11	*12*	*13*	14	*15*	16
1	02	163026	146799	166794	116227	139188	153849	117740	142667
	3	17	*18*	*19*	20	21	*22*	23	*24*
1	03	150172	136201	131706	152238	160820	132999	143980	127587
	4	*25*	26	27	*28*	29	*30*	31	32
1	04	120404	155357	148280	120105	153389	128056	151212	151241
	5	33	34	35	*36*	*37*	*38*	*39*	40
1	05	151347	150316	155497	135453	126640	119047	131650	162577
	6	41	42	43	*44*	45	46	47	*48*
1	06	167694	146046	143138	133354	163489	144113	161846	134956
	7	49	*50*	*51*	52	53	*54*	55	56
1	07	166754	126042	123752	154003	154452	136909	146602	154498
	8	57	58	59	*60*	*61*	62	63	*64*
1	08	143104	156416	144844	117903	126260	155789	140611	132597
	9	*65*	66	67	68	*69*	70	71	72
1	09	126666	158603	155020	154630	131846	150880	144715	145562
	10	73	*74*	*75*	76	*77*	*78*	*79*	80
1	10	149807	132131	111299	140000	134906	121384	119207	172935

TABLE 11
The Factor-Loading Calculation Sheets (See operation 4)

Sec 1	+31 1	+20 2	-05 (3)	+57 4	-16 (5)	+20 6	-28 (7)	-22 (8)	a.
1	156136	149989	138699	169916	132219	150699	127562	131447	B_1
	140000	140000	140000	140000	140000	140000	140000	140000	B_2
	16136	9989	-1301	29916	-7781	0699	-12438	-8553	B_3
	2766	2520	-1927	5127	-2275	1496	-4793	-4793	C
	18902	12509	-3228	35043	-10056	12195	-17231	-13346	D
Sec 2	+43 9	+18 10	+48 11	-48 (12)	-08 (13)	+26 14	-40 (15)	+09 16	a
1	163026	146799	166794	116227	139188	153849	117740	142667	B_1
	140000	140000	140000	140000	140000	140000	140000	140000	B_2
	23026	6799	26794	-23773	-0812	13849	-22260	2667	B_3
	3218	4140	2426	-2639	-4140	2039	-1734	3028	C
	26244	10939	29220	-26412	-4953	15888	-23994	5695	D
Sec 3	+21 17	-10 (18)	-16 (19)	+26 20	+37 21	-22 (22)	+13 23	-23 (24)	a
1	150172	136201	131706	152238	160820	132999	143980	127587	B_1
	140000	140000	140000	140000	140000	140000	140000	140000	B_2
	10172	-3799	-8294	12238	20820	-7001	3980	-12413	B_3
	2991	-2679	-1720	3920	1895	-6536	4017	-1748	C
	13163	-6478	-10014	16158	22715	-13537	7997	-14161	D

TABLE 11a

The Verification Table for Sign-Change Procedure

Table total B_1 rows	$=B_1t$	1142.3896
Correction factor, total B_2 rows	$=B_2t$	1120.0000
Table total B_3 rows	$=B_3t$	22.3896
Communalities total	$=Ct$	4.3772
Table total D rows	$=Dt$	26.7668
Sum of factor loadings		4.39
Calculated factor loading total $D_t \times 1/\sqrt{r_t} =$		4.41

TABLE 11b

Table Used to Calculate $1/r_t$

Table totals without regards for signs

Table total B_3 rows	30.5106
Communalities total	6.3000
Table total D rows	36.8106
$\sqrt{r_t} = \sqrt{36.8106} =$	6.067174
$1/\sqrt{r_t} =$	.16482

TABLE 12

Factor-Loading Tabulation Sheet

Unrotated factor loadings for each variable

Variable	1	2	3	4	5	6	7	8	9	10
1	−.03	+.42	+.35	+.10	+.07	+.27	+.21	+.14	−.10	+.31
2	+.34	+.22	+.27	+.16	+.19	−.29	+.10	+.10	−.15	+.20
3	+.21	−.12	−.18	+.17	+.11	−.13	−.10	+.11	−.08	−.05
4	+.23	−.27	−.42	+.19	−.40	−.46	−.30	−.21	−.25	+.57
5	−.03	+.06	+.24	−.02	+.13	−.22	+.05	+.03	−.20	−.16
6	−.13	−.03	+.15	−.29	+.08	+.05	+.17	+.01	−.07	+.20
7	−.10	−.26	+.19	+.17	+.08	−.32	+.10	+.22	−.21	−.28
8	+.09	−.28	+.17	+.12	+.15	−.26	+.14	+.20	−.16	−.22
Section Total	+.78	−.26	+.77	+.60	+.41	−1.36	+.37	+.60	−1.22	+.57
9	+.42	+.14	+.42	−.35	−.31	+.25	−.22	+.29	+.04	+.43
10	+.59	+.17	−.41	−.09	+.25	−.14	+.25	−.22	−.16	+.18
11	−.48	+.07	−.18	+.11	−.06	+.02	+.13	−.17	+.13	+.48
12	−.26	−.26	−.26	+.31	+.37	−.13	+.11	−.10	+.04	−.44
13	−.28	−.33	−.09	+.14	+.27	−.06	+.09	+.07	−.16	−.08
14	−.05	−.25	−.15	−.21	+.30	−.13	+.13	−.10	+.16	+.26
15	+.60	+.04	−.07	+.27	+.26	−.14	−.18	+.27	−.13	−.40
16	−.48	−.06	+.13	+.23	−.11	+.17	+.14	+.08	+.14	+.09
Section Total	+1.02	−.48	−.61	+.41	+.97	−.16	+.43	+.12	+.06	+.52

TABLE 12 (Continued)
Factor-Loading Tabulation Sheet
Unrotated factor loadings for each variable

Variable	1	2	3	4	5	6	7	8	9	10
17	+.29	—.12	+.21	—.21	+.14	+.20	—.11	+.09	—.03	+.21
18	—.36	+.08	—.36	—.38	—.07	—.18	+.15	+.29	—.13	—.10
19	+.73	+.26	—.39	—.21	+.08	—.16	+.28	+.12	+.02	—.16
20	—.53	+.03	+.08	—.04	—.13	—.15	—.12	+.12	—.06	+.26
21	—.27	+.10	—.05	+.10	—.14	—.02	—.13	—.12	—.19	+.37
22	+.44	+.07	—.11	—.17	+.17	+.15	—.16	—.17	+.31	—.22
23	+.58	+.20	+.10	+.18	+.34	—.19	—.07	—.10	—.07	+.13
24	—.73	—.28	+.36	+.13	+.04	+.15	—.27	—.07	—.07	—.23
Section Total	+.15	+.34	—.16	—.60	+.43	—.20	—.43	+.16	—.22	+.26
25	—.60	+.07	—.21	+.16	—.18	+.21	—.31	+.06	+.11	—.37
26	—.31	+.29	—.17	—.13	+.12	—.23	—.20	+.08	—.14	+.32
27	—.27	—.06	—.15	+.22	+.14	—.14	+.27	+.17	+.23	+.18
28	—.39	+.11	+.03	+.09	—.12	—.25	—.11	+.13	—.05	—.37
29	+.38	+.17	—.22	—.25	+.46	+.26	—.30	—.07	—.15	+.29
30	—.28	—.04	—.06	—.14	+.04	+.17	—.17	+.26	+.27	—.24
31	—.11	+.32	+.19	—.18	—.25	—.18	—.08	+.16	+.17	+.23
32	+.05	—.19	+.05	—.12	—.19	+.21	+.14	+.13	—.18	+.21
Section Total	—1.53	+.67	—.54	—.35	+.02	+.05	—.76	+.92	+.26	+.25
Variable	1	2	3	4	5	6	7	8	9	10
33	+.44	+.08	+.05	+.04	+.13	+.06	+.24	—.11	—.27	+.23
34	+.36	—.50	—.08	+.23	—.19	—.09	+.17	+.23	—.09	+.23
35	+.36	—.31	—.10	+.28	+.07	+.18	+.04	—.24	+.24	+.30
36	+.05	+.13	—.15	+.16	—.16	+.13	+.07	—.12	—.10	—.12
37	+.57	+.24	+.11	+.11	+.19	—.33	—.13	+.04	+.19	—.27
38	—.48	+.28	+.14	+.29	+.12	+.31	+.18	+.09	—.22	—.43
39	+.57	+.29	+.22	—.09	+.06	+.10	+.12	—.17	—.11	—.17
40	—.16	—.18	+.08	+.09	+.06	—.07	+.07	+.13	—.05	+.45
Section Total	+1.71	+.03	+.27	+1.01	+.28	+.29	+.76	—.15	—.41	+.22
41	+.21	—.24	+.22	+.16	+.05	—.12	+.24	+.12	+.28	+.50
42	+.66	+.43	—.35	—.28	—.13	—.12	+.39	+.14	+.22	+.14
43	—.41	—.26	+.22	—.09	+.18	+.27	—.25	—.19	—.22	+.09
44	+.48	+.34	+.10	+.06	+.05	—.20	—.31	—.32	+.14	—.16
45	—.30	—.13	—.26	—.51	+.22	+.25	+.36	+.12	—.24	+.44
46	—.19	+.18	—.11	—.03	+.04	+.03	+.26	+.08	+.27	+.11
47	—.15	—.04	+.09	—.11	+.15	—.11	—.10	—.07	+.11	+.40
48	+.50	—.22	+.35	+.38	+.11	—.07	—.07	—.17	+.21	—.12
Section Total	+.80	+.06	+.26	—.42	+.67	—.07	+.52	—.57	+.77	+1.40

TABLE 12 (Continued)
Factor-Loading Tabulation Sheet
Unrotated factor loadings for each variable

Variable	1	2	3	4	5	6	7	8	9	10
49	+.45	+.13	+.39	—.35	—.12	+.08	—.36	+.31	—.03	+.48
50	+.10	+.19	—.26	+.23	+.18	+.03	+.10	+.29	+.13	—.26
51	+.44	—.11	+.20	—.12	—.19	+.12	—.10	—.23	—.15	—.32
52	—.49	—.14	—.11	+.15	—.11	+.28	—.09	+.09	—.32	+.28
53	+.55	—.06	—.14	—.33	+.20	—.18	—.29	+.23	—.16	+.34
54	+.75	+.23	+.15	—.14	—.10	+.05	+.32	—.13	—.05	—.09
55	—.38	+.12	—.23	+.09	—.16	+.06	—.10	+.11	+.25	+.16
56	—.16	+.32	+.08	+.06	+.15	+.13	+.15	—.14	+.24	+.29
Section Total	+1.26	+.68	+.08	—.41	—.13	+.57	—.38	+.53	—.09	+.88
57	+.65	—.08	+.05	—.29	+.17	—.10	—.22	—.04	—.06	+.09
58	+.23	+.11	+.18	+.22	—.22	+.18	+.23	—.12	+.10	+.32
59	—.12	—.09	+.37	—.16	+.08	+.10	+.16	+.22	+.20	+.15
60	+.54	+.19	+.26	—.16	—.14	+.05	+.16	—.16	+.15	—.43
61	—.57	+.19	+.22	+.38	+.13	+.19	—.20	+.30	+.21	—.31
62	—.02	—.23	+.33	+.20	—.23	—.14	+.06	—.20	+.08	+.30
63	+.23	+.42	+.22	—.09	+.18	+.12	+.09	+.32	—.18	+.05
64	+.41	—.13	—.10	—.12	—.15	+.28	+.10	+.27	+.12	—.17
Section Total	+1.35	+.38	+1.53	—.02	—.18	+.88	+.38	+.59	+.62	.00
65	+.08	+.24	+.12	+.03	+.11	+.20	—.24	+.26	+.08	—.27
66	+.36	—.20	+.05	+.07	—.08	+.15	+.03	—.08	+.26	+.34
67	—.08	+.25	—.15	+.26	+.21	+.30	—.17	—.21	—.34	+.32
68	—.12	+.16	—.23	+.18	+.22	+.29	+.13	—.21	—.35	+.33
69	—.21	+.29	+.16	—.10	+.14	—.24	+.06	—.03	+.17	—.16
70	+.13	—.24	—.07	+.10	+.22	+.17	+.10	+.11	+.22	+.23
71	—.45	+.12	+.26	+.30	—.20	—.02	+.20	+.13	—.14	+.11
72	—.21	+.30	+.12	—.11	+.17	—.14	+.11	—.10	+.21	+.13
Section Total	—.50	+.92	+.26	+.73	+.79	+.71	+.22	—.13	+.11	+1.03
73	+.22	+.07	+.15	+.16	+.11	+.08	+.09	+.09	+.04	+.20
74	—.50	—.06	—.10	+.26	+.14	—.10	—.07	+.10	+.33	—.17
75	—.29	—.47	+.04	—.11	—.07	+.07	+.05	—.17	—.12	—.54
76	00	00	00	00	00	00	00	00	00	00
77	+.32	+.05	—.11	+.21	+.12	+.15	+.12	+.05	+.10	—.11
78	+.29	—.20	—.24	+.22	—.19	+.18	+.06	+.13	—.18	—.36
79	+.33	—.22	+.17	+.24	+.10	+.19	+.11	—.11	—.13	—.38
80	+.22	+.07	+.14	+.28	+.34	+.23	—.09	+.10	+10	+.62
Section Total	+.59	—.76	+.04	+1.26	+.55	+.80	+.27	+.19	+.14	—.74
Column Total	5.63	1.58	1.90	2.21	3.81	1.51	1.40	2.26	.02	4.39

PSYCHOMETRIKA—VOL. 10, NO. 2
JUNE, 1945

A NOTE ON RELIABILITY

HYMAN B. KAITZ
SOCIAL SECURITY BOARD
(on military leave.)

A formula for internal consistency reliability is developed within the framework of the analysis of variance. The test items are assumed to be homogeneous, but may have any weights. Data needed for computation are the student test scores, and the total number of items answered so as to have the same weight. It is shown that this formula reduces to the Kuder-Richardson (21) for item weights of one and zero. Some empirical validation is offered.

When both the distribution of total scores and the item count of a test are available, it is convenient to estimate the reliability by the use of Kuder-Richardson formula (20).* When only the distribution of total scores is available, a simple and adequate approximation to the reliability coefficient can be made with Kuder-Richardson formula (21). As these formulas were originally stated, they applied only to the case of a test consisting of items weighted one (right) or zero (wrong, or omit). The concept of K-R (20) was put on a broader foundation when it was shown that this formula could be stated in terms of analysis of variance components.† In addition, its derivation was given rigor by the method of estimation of error variance. Any weighting of items whatever could be used in the Hoyt formula. For specific methods of weighting test items, a number of variations of K-R (20) were offered by another author.‡ There is, however, no general formula similar to K-R (21) available for a system of item weighting other than one or zero. Such a formula, the conditions for its use, and some empirical results are reported below.

Development

Hoyt§ expressed in equation form the assumption that the score of an individual is the sum of four independent factors:

$$X_{is} = A + t_i + p_s + y_{is}, \qquad (1)$$

* Kuder, G. F. and Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika*, 1937, 2, 151-160.

† Hoyt, C. J. Test reliability estimated by the analysis of variance. *Psychometrika*, 1941, 6, 153-160.

‡ Dressel, P. L. Some remarks on the Kuder-Richardson reliability coefficient. *Psychometrika*, 1940, 5, 305-310.

§ *Op. cit.*

where

X_{is} = observed score of individual s on item i;
A = factor common to all individuals and all items;
t_i = factor associated with item i;
p_s = factor associated with individual s; and
y_{is} = error factor, normally distributed, with zero mean.

Using this assumption, he derived the best linear estimate of error variance in the sense of least squares, and the best estimate of the true variance among individuals as the difference between the variance estimate "among individuals" and the error variance estimate. The reliability is then written as the ratio of the true variance to the observed variance, thus:

$$r_{tt} = \frac{V_s - V_e}{V_s}, \tag{2}$$

where

V_s = variance estimate among individuals and
V_e = variance estimate of error (interaction).

If now, we make the assumption, far less exacting than the one above, that an individual's score is the sum of three factors, we may write

$$X_{is} = A + p_s + y_{is}, \tag{3}$$

where the factors have the same meaning as before, but we are assuming that all items have the same factor, which will therefore be included in A, the common factor. By following the same procedure as Hoyt, and minimizing the sum of squares, $\sum_i \sum_s y_{is}^2$, with respect to A and p_s, we shall find that the best estimate of error variance is what is commonly known as the variance "within individuals." Considering the sums of squares involved, we may separate the total into two components: the sum of squares due to differences among individuals (SS_s), and the sum of squares within individuals (SS_w). This is

$$SS_T = SS_s + SS_w. \tag{4}$$

The best estimate of error variance is

$$V_w = \frac{SS_w}{n_s(n_i - 1)}, \tag{5}$$

where n_s is the number of individuals and n_i is the number of items.

Our reliability formula may be expressed as

$$r_{tt} = \frac{V_s - V_w}{V_s}, \tag{6}$$

where V_s is the same as in (2), but V_w comes from (5). This formula may be put in terms of the same elements as the Kuder-Richardson formulas by a simple algebraic change. This is

$$r_{tt} = \frac{\sigma_s^2 - \frac{(n_s - 1)n_i}{(n_i n_s - 1)}\left[\frac{\sum X_{ij}^2}{n_s} - \frac{\bar{X}_s^2}{n_i}\right]}{\sigma_s^2} \cdot \frac{n_i n_s - 1}{n_s(n_i - 1)}, \tag{7}$$

where σ_s^2 is the sample variance of individual total test scores and $\bar{X}_s$ is the mean of the same scores.

If the test items have k possible weights, $w_i (i = 1, \cdots, k)$, and the total number of items, for all individuals, answered so as to have a weight of w_j is N_j (where $\sum_{j=1}^{K} N_j = n_i n_s$), then we may express (7) as

$$r_{tt} = \frac{\sigma_s^2 - \frac{(n_s - 1)n_i}{n_i n_s - 1}\left[\frac{\sum^{K} N_i w_i^2}{n_s} - \frac{\bar{X}_s^2}{n_i}\right]}{\sigma_s^2} \cdot \frac{n_i n_s - 1}{n_s(n_i - 1)}. \tag{8}$$

At this point we may compare this formulia with K-R (21). By substituting $N_1/n_i n_s = \bar{p}$, $N_2/n_i n_s = \bar{q}$, $w_1 = 1$, and $w_2 = 0$, we get

$$r_{tt} = \frac{\sigma_s^2 - \frac{(n_s - 1)n_i}{n_i n_s - 1}(n_i \bar{p}\bar{q})}{\sigma_s^2} \cdot \frac{n_i n_s - 1}{n_s(n_i - 1)}. \tag{9}$$

For relatively large $n_s - 1$, and consequently large $n_i n_s - 1$, we may substitute n_s and $n_i n_s$, respectively, with little loss of accuracy. Formula (9) then becomes

$$r_{tt} = \frac{\sigma_s^2 - n_i \bar{p}\bar{q}}{\sigma_s^2} \cdot \frac{n_i}{n_i - 1}, \tag{10}$$

which is K-R (21). The same substitution changes (8) to

$$r_{tt} = \frac{\sigma_s^2 - \left[\frac{\sum N_i w_i^2}{n_s} - \frac{\bar{X}_s^2}{n_i}\right]}{\sigma_s^2} \cdot \frac{n_i}{n_i - 1} \tag{11}$$

Use of Formula (11) and Empirical Validation

Formula (11) is convenient to use when the item counter on the I.B.M. Test Scoring Machine is not available, or is not desired in favor of a quicker estimate of reliability. If papers are to be scored by hand, it is only necessary to see that each paper has k values entered on it, each one representing the number of items answered so as to have the same weight, $w_i (i = 1, \cdots, k)$. The sum of these values on each sheet is the total number of items in the test. The weighted sum of values on each sheet is the total score for that individual. If papers are to be scored by machine, the same values can be obtained by running the papers through the scoring machine fewer times than is necessary for a split-half estimate of reliability. Whenever items are marked right or wrong, and omits are treated as wrongs so that only two weights are possible, K-R (21) or a slight modification of it (if weights are other than one or zero) will suffice for a reliability estimate. If omits are weighted differently from rights or wrongs, formula (11) can be easily used if each test paper has the number of rights and the number of wrongs listed. The total number of omits can then be found from

$$\sum \text{Omits} = n_i n_s - \sum \text{Rights} - \sum \text{Wrongs}. \qquad (12)$$

A common type of test with multiple weights for the individual items is the personality or attitude scale. In many of these tests each item has five choices expressing varying degrees of approval, with weights of 1, 2, 3, 4, and 5 attached to these choices. A sample item is:

The government should not meddle with private business.

Choices:	(1) Strongly disagree	(2) Disagree	(3) Undecided
Weights:	1	2	3

Choices:	(4) Agree	(5) Strongly agree
Weights:	4	5

Here each test paper need have no more than four values on it; however, the use of all five values will serve as a check on the scoring.

In a study a number of years ago,* there are available sufficient data for a comparison of the usual split-half reliability with Hoyt reliability and with reliability based on formula (11) in the present paper. The tests used by Rundquist and Sletto are personality scales

* Rundquist and Sletto. Personality in the depression. Minneapolis: Univ. Minn. Press, 1936.

in the areas of morale, inferiority, family, law, economics and education. Each of the scales has 22 items, and each item has five choices as in the illustration above. The necessary data are taken from table 19, page 45, and table 44, page 85. Reliabilities were computed separately for two groups, one of 500 males and one of 500 females. The results are given in the table below.

Comparison of Three Reliability Coefficients

Scale	500 Males			500 Females		
	S-B	Hoyt	Formula (11)	S-B	Hoyt	Formula (11)
Morale	.79	.77	.73	.81	.81	.78
Inferiority	.78	.80	.77	.82	.81	.76
Family	.83	.78	.74	.87	.83	.81
Law	.84	.82	.78	.82	.83	.77
Economic Conservatism	.85	.82	.80	.82	.77	.74
Education	.82	.84	.81	.83	.84	.80

PSYCHOMETRIKA—VOL. 10, NO. 2
JUNE, 1945

FACTORIAL DESIGN AND COVARIANCE IN THE STUDY OF INDIVIDUAL EDUCATIONAL DEVELOPMENT*

PALMER O. JOHNSON AND FEI TSAO
UNIVERSITY OF MINNESOTA

This is the report of the application of the principles of factorial design to an investigation of individual educational development. The specific type of factorial design formulated was a $2 \times 3 \times 3 \times 3$ arrangement, that is, the effect of sex, grade location, scholastic standing, and individual order, singly and in all possible combinations was studied in relation to educational development as measured by the *Iowa Tests of Educational Development*. An application of the covariance method was introduced which resulted in increased precision of this type of experimental design by significantly reducing experimental error. The two concomitant measures used to increase the sensitiveness of the experiment were initial status of individual development and mental age. Without these statistical controls all main effects and two first-order interactions would have been accepted as significant. With their use only sex (doubtful), scholastic standing, and individual order demonstrated significant effects. The chief beauty of the analysis of variance and covariance as an integral part of a self-contained experiment is demonstrated in the complete single analysis of the data. The statistical utilization of the experimental results has also been developed for purposes of estimation and prediction. The mathematical statistician is being continuously required to develop and analyze experimental designs of increasing complexity since the introduction of the analysis of variance and covariance. The mathematical formulation and solution of the problem of this investigation is carried out. The methods illustrated and explained in this study, and modifications and extensions of them are capable of very wide application. The general principles can be used to various degrees and in a number of ways.

An earlier paper† has described in detail the use of factorial design in an experiment in psychology. It is proposed to report here the application of the principles of factorial design and the appropriate statistical analysis to an investigation of individual educational development. While the experimental results are themselves important, we wish to develop and illustrate the procedures which supply a sufficient technique for collecting the primary data in such a manner that valid conclusions may be educed from them, and for extracting the relevant information contained in the data in the most efficient way.

* For the research grant to finance this study, grateful acknowledgment is given to the Graduate School, the University of Minnesota.

† Johnson, Palmer O. and Tsao, Fei. Factorial design in the determination of differential limen values. *Psychometrika*, 1944, 9, 107-145.

The procedures are illustrated by original data which were collected in the University of Minnesota high school. The principal educational factors chosen for study were sex, scholastic standing, and individual order of educational development. Other factors could have been included, but the above were considered particularly worthy of study and sufficient in number so as not to risk losing the attention of the reader unfamiliar with the technique in a maze of unavoidable details, as might happen were numerous factors included.

It should be emphasized that the present work attempts to make available to the scientific reader the application of principles from the point of view of their utilization in the construction of more efficient and comprehensive experimental designs and clearer, more concise, and more efficient statistical analysis of results. The situation in this study is one of very wide occurrence. The traditional method for the study of the educational factors considered here would require a series of experiments each designed to test a single factor with arbitrary control of the other factors.* In modern designs the factors are tested simultaneously, which makes it possible to explore the interactions as well as the main effects. The specific design developed for this study was a $2 \times 3 \times 3 \times 3$ factorial type, the two sexes, three grades, three scholastic standings, and three individual orders. The type of factorial design is of the kind where absolute replication is dispensed with and hidden replication is involved.† This is the type desirable when large numbers of combinations are tested simultaneously without repeated use of each combination. Most of the higher-order interactions are assumed to be statistically insignificant. All of the independent comparisons contained in the experiment are allocated to the factors tested and their interactions. Since there is no independent comparison ascribable to pure error, the highest-order interactions are used as a basis for measuring the precision of the main comparisons.

A distinctive application to factorial design in this study is that of the analysis of covariance, which increases the precision of the experiment by use of the internal evidence of a self-contained experiment to reduce experimental error. By the precision of an experiment is meant the quantity of information it yields relevant to the hypothesis under test. In this study the concomitant measurements of two variables were used to increase the precision of the experimental comparisons. The statistical utilization of the findings has also been developed for purposes of estimation and prediction.

* *Ibid.*, p. 108.

† Fisher, R. A. The design of experiments. London: Oliver and Boyd, 1944, pp. 111-113.

The report is presented in two parts. Part I describes the design and analysis of the investigation. The mechanization of the calculations required by the research worker for the analysis and interpretation of the data from modern experimental designs is a distinctive feature of Fisher's contribution to the statistical craft. However, for the mathematical reader interested in the statistical model, Part II, which develops the mathematical formulation and solution of the problem, is presented.

PART I

The subjects in this investigation were students in the University of Minnesota High School. The *Iowa Tests of Educational Development** were administered to all students in the ninth, tenth, and eleventh grades in September, 1942, and again in September, 1943. The composite scores were used. This score is comprised of the scores on each of the following nine tests:

1. Understanding of Basic Social Concepts
2. Background in the Natural Sciences
3. Correctness in Writing
4. Ability to Do Quantitative Thinking
5. Ability to Interpret Reading Materials in the Social Studies
6. Ability to Interpret Reading Materials in the Natural Sciences
7. Ability to Interpret Literary Materials
8. General Vocabulary
9. Uses of Sources of Information

Standard scores are used in terms of a standard scale from 0 to 30 with a mean of 15, determined from the combined grades.

From the 168 students who had taken the two batteries of examinations, 54 were selected, 18 each from the tenth, eleventh, and twelfth grades.† The 18 individuals in the tenth grade were all of chronological age 15; those in the eleventh grade, 16; and those in the twelfth grade, 17. All mental ages were based on the scores from the *Terman Group Test of Mental Ability*. Since the intelligence tests were not administered to all students on the same date, mental ages for all were calculated as of the same date, i.e., May 21, 1939. Stu-

* College of Education, State University of Iowa, Iowa City, Iowa, 1942.

† We designed the investigation so as to secure equal representation in the sub-classes. While this equality is not an essential condition, it is highly desirable in order to avoid rather complicated mathematical formulation thus leading to a more laborious statistical analysis.

dents were assigned to scholastic groups on the basis of their honor-point ratios as follows:

(1) Good —honor-point ratios equal to or greater than 2
(2) Average—honor-point ratios between 1 and 2
(3) Poor —honor-point ratios less than 1.

Individual order of educational development was based upon the scores made on the second administration of the battery of examinations described above. The individual who attained the highest score was given rank 1, the individual next highest, rank 2, and the lowest score, rank 3. Where individuals were tied for rank the position was randomly determined.

All samples were selected at random by the use of Fisher and

TABLE 1*
Classification of the Raw Data of This Study

Sex	Scholastic Standing	Individual Order	Grade 10			Grade 11			Grade 12		
			Final	Initial	M.A.	Final	Initial	M.A.	Final	Initial	M.A.
Male	Good	1	30	28	45	26	22	62	29	25	60
		2	25	22	58	26	21	57	29	24	88
		3	22	19	46	24	21	65	22	19	64
	Average	1	26	22	56	24	25	54	23	21	64
		2	17	14	19	23	18	55	20	17	47
		3	14	14	29	15	13	24	19	17	75
	Poor	1	18	18	34	18	17	40	17	16	29
		2	17	14	17	16	13	24	15	15	38
		3	12	9	19	13	12	23	14	12	28
Female	Good	1	21	16	44	26	22	60	33	29	94
		2	21	21	44	25	22	57	29	29	89
		3	19	17	6	23	19	52	25	22	78
	Average	1	20	18	38	22	19	54	23	21	50
		2	18	16	27	21	19	54	18	19	57
		3	14	14	18	17	16	52	17	17	43
	Poor	1	14	9	18	19	17	40	15	13	36
		2	12	7	18	15	12	28	15	14	35
		3	9	7	5	13	12	48	10	9	14

* The chronological age of every individual in grade 10 is 15 years; grade 11, 16 years; grade 12, 17 years. A hundred has been subtracted from Mental Age (in terms of months). Moreover, it has been assumed the mental test was administered on the 21st of May in 1939.

Yates' Random Sampling Numbers.* The primary data are shown grouped into the several sub-classes in Table 1. The sums of squares and of products for all sources of variation are recorded in Table 2.

TABLE 2
Sums of Squares and of Products for All Variations*

Source of Variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$
Sex × Grade × Scholastic × Individual	8	16	26	690	13	21	—11
Sex × Grade × Scholastic	4	25	35	942	29	130	135
Sex × Grade × Individual	4	18	22	659	19	41	29
Sex × Scholastic × Individual	4	5	9	515	2	10	19
Grade × Scholastic × Individual	8	33	28	984	27	112	118
Sex × Grade	2	32	53	638	40	138	160
Sex × Scholastic	2	2	20	11	6	—4	—14
Sex × Individual	2	3	32	119	7	—14	—10
Grade × Scholastic	4	37	29	1299	27	213	138
Grade × Individual	4	5	7	445	0	22	38
Scholastic × Individual	4	10	17	373	9	28	76
Sex	1	30	19	69	24	45	36
Grade	2	62	83	5837	70	594	697
Scholastic	2	1037	824	9228	920	3073	2757
Individual	2	292	220	1022	253	545	471
Total	53	1607	1424	22831	1446	4954	4639

* Where y = deviation score from the grand mean of final
x_1 = deviation score from the grand mean of initial
x_2 = deviation score from the grand mean of M.A.
This will be true for all the following tables.

The Test of Significance of Interactions and Main Effects

We proceeded first to test the null hypothesis with respect to the interactions. Since no measure of pure error was available, the triple interaction, sex × grade × scholastic standing × individual, was used as error against which was compared each of the four second-order and each of the six first-order interactions† (Table 3). The null hypothesis was accepted for all interactions except two, sex × grade, and grade × scholastic standing.

All of the interactions shown to be insignificant were pooled and constituted a mean square of 2.93, with 40 degrees of freedom for

* Fisher, R. A., and Yates, F. Statistical tables. London: Oliver and Boyd, 1938, pp. 82-87.

† It is customary to call the interaction involving two factors, an interaction of first order; to call the interaction involving three factors, an interaction of second order; etc. Sometimes for the sake of brevity, the terms single, double, and triple, interactions are used to denote the interactions of first order, second order, and third order, respectively. For instance, we may call the interaction sex × grade × scholastic × individual, an interaction of third order. Or we may call it a triple interaction. Particularly in this study, this interaction may be regarded as of highest order.

TABLE 3

Test of Significance of Interactions
(Considering final score only)

Source of Variation	d.f.	Sum of Squares	Mean Square	F*	Hypothesis† Tested
Sex × Grade × Scholastic × Individual (or Error)	8	16	2.00	—	—
Sex × Grade × Scholastic	4	25	6.25	3.12	Accepted
Sex × Grade × Individual	4	18	4.50	2.25	Accepted
Sex × Scholastic × Individual	4	5	1.25	—	Accepted
Grade × Scholastic × Individual	8	33	4.13	2.06	Accepted
Sex × Grade	2	32	16.00	8.00	Remain in doubt
Sex × Scholastic	2	2	1.00	—	Accepted
Sex × Individual	2	3	1.50	—	Accepted
Grade × Scholastic	4	37	9.25	4.63	Remain in doubt
Grade × Individual	4	5	1.25	—	Accepted
Scholastic × Individual	4	10	2.50	1.25	Accepted

* Where $F = \frac{\text{Mean square of a tested variation}}{\text{Mean square of error or residual}}$

By referring to Snedecor's tables of F (see Snedecor, G. W., Statistical methods applied to experiments in agriculture and biology. Iowa: Collegiate Press, 1938, pp. 174-77), we may use the following four rules in testing the hypothesis: (a) reject the hypothesis tested, if the calculated value of F is greater than the 1% point given in the tables; (b) accept the hypothesis tested, if the calculated value of F is less than the 5% point given in the tables; (c) remain in doubt, if the calculated value of F lies between the 5% and 1% points given in the tables; (d) in the event that the mean square of error or residual is greater than the mean square of a tested variation, then we simply accept the hypothesis without calculating the value of F.

These conventions are always true throughout the following tables.

† The hypothesis tested is a null hypothesis concerning the variation in the same row. For instance, the hypothesis regarding sex × grade is that there is no significant interaction between sex and grade.

TABLE 4

Complete Analysis of Variance
(Considering final score only)

Source of Variation	d.f.	Sum of Squares	Mean Square	F	Hypothesis* Tested
Residual	40	117	2.93	—	—
Sex × Grade	2	32	16.00	5.46	Rejected
Grade × Scholastic	4	37	9.25	3.16	Remain in doubt
Sex	1	30	30.00	10.24	Rejected
Grade	2	62	31.00	10.58	Rejected
Scholastic	2	1037	518.50	176.96	Rejected
Individual	2	292	146.00	49.83	Rejected
Total	53	1607			

* The hypothesis tested is a null hypothesis regarding the variation in the same row. For instance, the hypothesis regarding grade is that there is no significant difference between grade-means.

TABLE 5

Test of Significance of Interactions
(Partialling out the effect of initial score)*

Source of Variation	d.f.	Σy^2	$\Sigma x_1 y$	Σx_1^2	Adjusted Σy^2	Reduced d.f.	Reduced S.S.	Reduced M.S.	Reduced F.	Reduced Hypothesis†
Sex × Grade × Scholastic × Individual (or Error)	8	16	13	26	9.50	7	9.50	1.36	—	—
Sex × Grade × Scholastic	4	25	29	35	4.75	4	2.58	.65	—	Accepted
Sex × Grade × Individual	4	18	19	22	4.50	4	3.17	.79	—	Accepted
Sex × Scholastic × Individual	4	5	2	9	5.25	4	5.07	1.27	—	Accepted
Grade × Scholastic × Individual	8	33	27	28	13.00	8	9.87	1.23	—	Accepted
Sex × Grade	2	32	40	53	5.25	2	2.94	1.47	1.08	Accepted
Sex × Scholastic	2	2	6	20	1.00	2	.65	.33	—	Accepted
Sex × Individual	2	3	7	32	4.00	2	2.60	1.30	—	Accepted
Grade × Scholastic	4	37	27	29	17.25	4	14.41	3.60	2.65	Accepted
Grade × Individual	4	5	0	7	6.75	4	6.38	1.60	1.18	Accepted
Scholastic × Individual	4	10	9	17	5.25	4	5.24	1.31	—	Accepted
Coefficients for adjusting Σy^2		1	—1.00	.25						

* Where $b = .50$

† The hypothesis tested is a null hypothesis concerning the variation in the same row. For instance, the hypothesis regarding Scholastic × Individual is that there is no significant interaction between Scholastic Standing and Individual Order when the effect of initial scores has been partialled out.

testing the significance of the main effects and the doubtful interactions. The complete analysis of variance of the final scores is shown in Table 4. All of the main effects, sex, grade, scholastic standing, and individual order were found to be significant. One of the two previous doubtful interactions, sex × grade, became significant, the other, grade × scholastic standing, remained doubtful.

Increasing the Precision by Concomitant Measurements

Corrections were introduced into the comparisons, which made allowances for the variations in the initial scores on the examinations and in mental age scores. The first factor to be controlled statistically was the initial score.

Influence of initial score. The necessary calculations for partialling out the effect of initial scores are given in Table 5. The theoretical basis developing the mathematical formulas are given in Part II and procedures necessary for the calculation of the sum of squares and sum of products for each source of variation are summarized in Table 1, Part II. The test of significance of the interactions is given in Table 5. None of the interactions was found to be significant after the influence of initial scores had been allowed for. It was noted in Table 3 that the interactions, sex × grade and grade × scholastic standing, were in the region of doubt.

The complete analysis of variance and covariance of the final scores, partialling out the effect of initial scores, is found in Table 6. It is to be noted that since none of the interactions was significant, 46 degrees of freedom, reduced to 45 after adjustment, were available

TABLE 6

Complete Analysis of Variance and Covariance
(Partialling out the effect of initial score)*

Source of Variation	d.f.	Σy^2	$\Sigma x_1 y$	Σx_1^2	Adjusted Σy^2	Reduced d.f.	S.S.	M.S.	F	Hypothesis‡
Residual	46	186	179	278	70.86	45	70.86	1.57	—	—
Sex	1	30	24	19	7.07	1	6.39	6.39	4.07†	Remain in doubt
Grade	2	62	70	83	6.43	2	6.75	3.38	2.15	Accepted
Scholastic	2	1037	920	824	197.24	2	66.81	33.41	21.28	Rejected
Individual	2	292	253	220	58.36	2	35.97	17.99	11.46	Rejected
Total	53	1607	1446	1424	339.96					
Coefficients for Adjusting Σy^2		1	—1.28	.41						

* Where $b = .64$

† Near the borderline of 5% significance and insignificance.

‡ The hypothesis tested is a null hypothesis concerning the variation in the same row. For instance, the hypothesis regarding sex is that there is no significant difference between sex-means when the effect of initial scores has been partialled out.

TABLE 7
Test of Significance of Interactions
(Partialling out the effects of both initial score and M.A.)*

Source of variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Adjusted Σy^2	Reduced d.f.	S.S.	M.S.	F.	Hypothesis†
Sex × Grade × Scholastic × Individual (or Error)	8	16	26	690	13	21	—11	8.47	6	8.47	1.41	—	—
Sex × Grade × Scholastic	4	25	35	942	29	130	135	1.13	4	.48	.12	—	Accepted
Sex × Grade × Individual	4	18	22	659	19	41	29	3.22	4	2.33	.58	—	Accepted
Sex × Scholastic × Individual	4	5	9	515	2	10	19	6.09	4	5.47	1.37	—	Accepted
Grade × Scholastic × Individual	8	33	28	984	27	112	118	10.11	8	8.92	1.12	—	Accepted
Sex × Grade	2	32	53	638	40	138	160	1.49	2	.62	.31	—	Accepted
Sex × Scholastic	2	2	20	11	6	—4	—14	.90	2	.60	.30	—	Accepted
Sex × Individual	2	3	32	119	7	—14	—10	5.16	2	3.38	1.19	—	Accepted
Grade × Scholastic	4	37	29	1299	27	213	138	7.83	4	3.62	.91	—	Accepted
Grade × Individual	4	5	7	445	0	22	38	7.35	4	6.47	1.62	1.15	Accepted
Scholastic × Individual	4	10	17	373	9	28	76	6.67	4	6.02	1.51	1.07	Accepted
Coefficients for Adjusting Σy^2		1	.2666	.0015	—1.0328	—.0774	.0400						

* Where $b_1 = .5164$; $b_2 = .0387$.
† The hypothesis tested is a null hypothesis concerning the variation in the same row. For instance, the hypothesis regarding Sex × Grade × Scholastic is that there is no significant interaction between sex, grade and scholastic standing when the effects of both initial score and mental ages have been partialled out.

for the exact tests of significance.* From the tests of significance it was found that scholastic standing and individual order were significant and that sex differences were doubtful. The increased precision in the application of covariance is observed in comparing the results with those reported in Table 4 where five main effects and one interaction were found to be significant and an additional interaction was in doubt.

Influence of initial score and mental age. We have now to consider the possible gain in precision which might result from eliminating, or making allowance for, two variables simultaneously. The two concomitant measurements were initial score and mental age. The application of the covariance method to two independent variables involves the calculation of the sums of squares of the dependent variable and of each of the two independent variables, and the sums of products of each with the dependent variate to be adjusted, and with each other (see Table 7).† The calculations for the interactions of all orders are found in Table 7, and the significance of each source of variation has been determined after partialling out the effects of initial score and mental age. Since the adjustments for the two concomitant variates have been obtained from the error term, two degrees of freedom ascribed to error have been used in evaluating it. The reduced sum of squares ascribed to error is divided by the degrees of freedom to obtain the mean square (1.41) appropriate to testing the significance of the remaining interactions. No significant interaction was found. There became available then 46 degrees of freedom before adjustment, reduced to 44 after adjustment, to test the significance of the main effects.

The complete analysis of variance and covariance of the final scores, partialling out the joint effect of initial score and mental age, is presented in Table 8. A comparison may be made by referring again to Table 6 between the mean adjusted $[\sum y^2]$'s after partialling out the effect of initial score and mental age and those after partialling out the effect of initial score only. The changes brought about by considering the joint effect of initial score and mental age did not alter the conclusions drawn when only the effect of initial score was eliminated.‡ The final analysis, then, which has utilized all the evi-

* See pp. 159-60, Part II, for explanation.

† See pp. 160-162, Part II, for the mathematical formulation of the procedures employed.

‡ By ordinary methods, we obtain the following three zero-order correlations from the scores of all the individuals:

$$r_{1y} = .96\,, \qquad r_{2y} = .82\,, \qquad r_{12} = .81\,,$$

where y denotes the final score; 1 denotes the initial score; and 2 denotes the mental age score. The first-order correlations with the third variable partialled

TABLE 8

Complete Analysis of Variance and Covariance
(Partialling out the effects of both initial score and M.A.)*

Source of Variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Adjusted Σy^2	Reduced d.f.	Reduced S.S.	Reduced M.S.	Reduced F.	Reduced Hypothesis†
Residual	46	186	278	6675	179	697	678	57.21	44	57.21	1.30	—	—
Sex	1	30	19	69	24	45	36	7.70	1	7.21	7.21	5.55	Remain in doubt
Grade	2	62	83	5837	70	594	697	3.25	2	2.89	1.45	1.12	Accepted
Scholastic	2	1037	824	9228	920	3073	2757	159.28	2	46.86	23.43	18.02	Rejected
Individual	2	292	220	1022	253	545	471	60.52	2	34.14	17.07	13.13	Rejected
Total	53	1607	1424	22831	1446	4954	4639	287.96					
Coefficients for Adjusting Σy^2		1	.2677	.0027	—1.0348	—.1038	.0536						

* Where $b_1 = .5174$; $b_2 = .0519$.

† The hypothesis tested is a null hypothesis concerning the variation in the same row. For instance, the hypothesis regarding grade is that there is no significant difference between grade-means when the effects of both initial score and mental ages have been partialled out.

dence of the relevant data, led to the conclusion that there was a significant difference between the mean final scores of scholastic standings and of the individual orders when adjustments were made for the differences in initial scores and in mental ages. Further observations would appear to be necessary to establish whether or not a highly significant sex difference existed. The difference was significant at the 5 per cent level.

Problems in Prediction and Estimation

In the previous section it was established that a significant difference existed among the means of final scores of the three scholastic standings and among the mean final scores of the individual orders. In this section we present the solution of the problems of estimating final scores from the scholastic standings and from the individual orders. There are two kinds of final scores, the observed and the predicted. Therefore, two kinds of regression equations have been developed, one of each for the scholastic standings and for the individual orders for the estimation of (1) observed final scores and (2) predicted final scores.* In other words, if one person has been classified in a certain individual order or in a certain scholastic standing what value of the observed and the predicted final score can be most probably got without having any other information? These scores can be obtained from the regression equations developed.†

The estimation from scholastic standings. We present first the equations for the estimation of observed final scores from the different scholastic standings. First, it is necessary to obtain the predicted final scores from initial scores and mental ages for each scholastic standing. Here the multiple regression equation is

$$\tilde{y} = b_1\bar{x}_1 + b_2\bar{x}_2 = .5174\bar{x}_1 + .0519\bar{x}_2 . \qquad (1)‡$$

* In the equations that follow we should like to state here the distinction in notation used: ~ is used as the symbol of prediction and ∧ of estimation.

† Errors of estimation can be calculated in the usual manner.

‡ See page 160, Part II for the development of this equation. The values of b_1 and b_2 were calculated from the residual component (see Table 8).

---

out are as follows:

$$r_{1y\cdot 2} = .87; \qquad r_{2y\cdot 1} = .24; \qquad r_{12\cdot y} = .19 .$$

This shows that if we regard the final score as a dependent variable, then the effect of the initial score is much more significant than that of the mental age. So in our case, by removing only the effect of the initial score we can largely increase the precision of our results. However, it does not necessarily mean that we need not remove the effect of the mental age. So far as there is some relationship between the mental age and the final score, it will always be advisable to use the technique of analysis of covariance to increase the precision of our results even if the increase is slight.

The necessary calculations are recorded in Table 9.

TABLE 9

Final Scores and Predicted Final Scores from Initial Scores and M.A. for Each Scholastic Standing*

(Using $\tilde{y} = b_1 \bar{x}_1 + b_2 \bar{x}_2 = .5174\ \bar{x}_1 + .0519\ \bar{x}_2$)

Scholastic	$\bar{Y}$	$\bar{X}_1$	$\bar{X}_2$	$\bar{y}$	$\bar{x}_1$	$\bar{x}_2$	$\tilde{\bar{y}}$	$\tilde{\bar{Y}}$
Good	25.2778	22.1111	59.3888	5.5000	4.6296	15.3333	3.1912	22.9690
Average	19.5000	17.7778	45.3333	−.2778	.2963	1.2778	.2196	19.9974
Poor	14.5556	12.5556	27.4444	−5.2222	−4.9259	−16.6111	−3.4108	16.3670
Total	19.7778	17.4815	44.0555					

* Where $\bar{Y}$ = mean final score, for instance, 25.2778 is the mean final score of all 18 individuals who have good standings.

$\bar{X}_1$ = mean initial score

$\bar{X}_2$ = mean mental age

$\bar{y}$ = mean deviation score from the grand mean of final score

$\bar{x}_1$ = mean deviation score from the grand mean of initial score

$\bar{x}_2$ = mean deviation score from the grand mean of mental age

$\tilde{\bar{y}}$ = predicted mean deviation score from the grand mean of final score

$\tilde{\bar{Y}}$ = predicted mean final score

It follows that $\tilde{\bar{y}} = \tilde{\bar{Y}} - 19.7778$,

$\tilde{\bar{Y}} = \tilde{\bar{y}} + 19.7778$.

From the mean final scores, $[\bar{Y}_2]$'s, of different scholastic standings we can obtain the estimated observed final scores, $[\dot{Y}]$'s from scholastic standing by the following regression equations:†

$$\text{Linear form: } \dot{Y} = 19.7778 + 5.3611x\,, \tag{2}$$

$$\text{Quadratic form: } \dot{Y} = 19.5000 + 5.3611x + .4167x^2\,, \tag{3}$$

where $x = 1\,,\ 0\,,\ -1$, which indicates good, average, and poor scholastic standings, respectively.

From Table 10, it is noted that the variation due to scholastic

† The details of the method of calculating the regression equations are not described here. As we wish to use the analysis of variance for testing the goodness of fit, it is convenient to use orthogonal polynomials. The interested reader is recommended to refer to the following sources:

Goulden, C. H. Methods of statistical analysis. New York: John Wiley, 1939, pp. 219-246.

Anderson, R. L., and Houseman, E. E. Tables of orthogonal polynomial values extended to $N = 104$. Agricultural Experiment Station, Iowa State College of Agriculture and Mechanics Arts, Research Bulletin 297, 1942, pp. 595-606.

TABLE 10

Components of Variation Due to Scholastic Standing
(Partialling out the effects of both initial score and M.A.)

Source of Variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Adjusted Σy^2	Reduced d.f.	S.S.	M.S.	F.	Hypothesis*
Linear	1	1035	822	9184	922	3083	2747	152.98	1	38.09	38.09	29.30	Rejected
Quadratic	1	2	2	44	−2	−10	10	6.30	1	6.28	6.28	4.83	Remain in doubt
Scholastic Standing	2	1037	824	9228	920	3073	2757	159.28					

* The hypothesis tested is a null hypothesis concerning the variation in the same row by using residual variance in Table 8 as a basis. For instance, the hypothesis concerning the linear component is that there is no significant linear relationship between final scores and scholastic standing when the effects of both initial score and mental ages have been removed.

standing is quadratically significant at the 5% level. It was therefore necessary to develop equation (3) of quadratic form. From this equation it is possible to estimate the observed final score from scholastic standing.

Likewise, from the mean predicted final scores, $[\bar{\bar{Y}}]$'s of different scholastic standings, we can obtain the regression equations for the estimation of predicted final scores, $[\hat{\hat{y}}]$'s, from different scholastic standings

$$\text{Linear form: } \hat{\hat{Y}} = 19.7778 + 3.3010x, \tag{4}$$

$$\text{Quadratic form: } \hat{\hat{Y}} = 19.9974 + 3.3010x - .3294x^2. \tag{5}$$

From equation (5) one can estimate the predicted final scores from scholastic standing.

The summary of the findings for variation due to scholastic standing is given in Table 11. The arithmetical means on final scores for good, average, and poor scholastic standings, were 25.3, 19.5, and 14.6, respectively. In Table 11, $\xi'_1(s)$ and $\xi'_2(s)$ denote the linear and the quadratic polynomial coefficients for each scholastic standing, respectively. These coefficients can be found in Fisher and Yates' Table 23.*

The sums of squares and of products for the two components of variation in Table 10 can be calculated from the figures given in Table 11. For instance, to obtain $\sum y^2$ linear (see Table 10), we have (i) to

TABLE 11

Summary of Results for Variation Due to Scholastic Standing

Scholastic Standing		Y	X_1	X_2	$\xi'_1(s)$	$\xi'_2(s)$
Good	Sums of Scores	455	398	1069	+1	+1
	Mean	25.2778	22.1111	59.3888		
Average	Sums of Scores	351	320	816	0	—2
	Mean	19.5000	17.7778	45.3333		
Poor	Sums of Scores	262	226	494	—1	+1
	Mean	14.5556	12.5556	27.4444		

* Fisher, R. A., and Yates, F. Statistical tables. London: Oliver and Boyd, 1938, pp. 54-59.

multiply each of three sums of Y scores (or mean Y scores) by the value $\xi'_1(s)$ in the same row as itself; (ii) add up the three products, (iii) square the sum; (iv) divide by 18, which is the number of individuals in each scholastic standing [if, in (i) we use the mean Y scores, then we multiply by 18 in this step]; and (v) divide by the sum of squares of $\xi'_1(s)$, namely, 2. Following these steps, we have

$$\frac{[445\ (+1) + 351\ (0) + 262\ (-1)]^2}{18\ (2)} = 1{,}035$$

or

$$\frac{[25.2778\ (+1) + 19.5000\ (0) + 14.5556\ (-1)]^2}{2} = 1{,}035\,.$$

As another example, to obtain $\sum x_1 y$ quadratic (see Table 10), we have (i) to multiply each of the three sums of Y scores (or mean Y scores) by the value of $\xi'_2(s)$ in the same row as itself and add these three products, (ii) multiply each of the three sums of X_1 scores (or mean X_1 scores) by the corresponding value of $\xi'_2(s)$ and add up the three products; (iii) find the products of the above two sums of products; (iv) divide by 18 [if in (i) we use the mean scores, then we multiply by 18]; and (v) divide by the sum of squares of $\xi'_2(s)$, namely, 6. Following these steps, we have

$$\{[455\ (+1) + 351\ (-2) + 262\ (+1)][39\ (+1) + 320\ (-2) + 226\ (+1)]\}/18\ (6) = -2$$

or

$$\{[25.2778\ (+1) + 19.5000\ (-2) + 14.5556\ (+1)][22.1111(+1) + 17.778\ (-2) + 12.5556\ (+1)]\ (18)\}/6 = -2\,.$$

The estimation from individual order. We now present the equation for the estimation of observed final scores from different individual orders. In Table 12 are tabulated the final scores, and predicted final scores from initial scores and mental ages for each individual order. The following regression equations were established for estimating the observed final scores from individual orders:

$$\hat{Y} = 19.7778 + 2.8334\,x\,, \qquad (6)$$

where $x = 1\,,\ 0\,,\ -1$, which denotes first, second, and third order, respectively. Only the linear component of variation due to individual order was found to be significant (see Table 13).

TABLE 12
Final Scores and Predicted Final Scores from Initial Scores and M.A. for Each Individual Order*

(Using $\tilde{\bar{y}} = .5174\ \bar{x}_1 + .0519\ \bar{x}_2$)

Individual	$\bar{Y}$	$\bar{X}_1$	$\bar{X}_2$	$\bar{y}$	$\bar{x}_1$	$\bar{x}_2$	$\tilde{\bar{y}}$	$\tilde{\bar{Y}}$
1	22.4445	19.8889	48.7777	2.6667	2.4074	4.7222	1.4907	21.2685
2	20.1111	17.6111	45.1111	.3333	.1296	1.0556	.1218	19.8996
3	16.7778	14.9445	38.2777	−3.0000	−2.5370	−5.7778	−1.6125	18.1653
Total	19.7778	17.4815	44.0555					

* Where all the notations are the same as in Table 9.

Finally, we present the equation for the estimation of predicted final scores from the different individual orders. Following the same procedures as before we obtained the linear equation

$$\hat{\tilde{Y}} = 19.7778 + 1.5516\,x\,, \tag{7}$$

where x has the same meaning as in (6).

The summary of the findings for variation due to individual order is given in Table 14. The mean final scores were 22.4, 20.1, and 16.8 for individual orders 1, 2, and 3, respectively. Their respective means on the initial examinations were 19.9, 17.6, and 14.9, and of the mental ages in months, 148.8, 145.1, and 138.3. In Table 14 $\xi'_1(i)$ and $\xi'_2(i)$ denote the linear and the quadratic polynomial coefficients for each individual order, respectively. The sums of squares and of products for the two components in Table 13 can be calculated from the figures given in Table 14 in the same way as indicated in the previous section.

Summary and Conclusions

1. The principles of factorial design were applied in this investigation of individual educational development. The specific type of factorial design used was a $2 \times 3 \times 3 \times 3$ arrangement, that is, the two sexes, three grades, three scholastic standings, and three individual orders.
2. The appropriate statistical analysis for this type of design is the analysis of variance. An application of the covarience method was introduced into the method of analysis to control two sources of variation, the mental age and the initial scores of the experimental subjects. In this way, the precision of the experimental comparisons was increased by reducing the experimental error.

TABLE 13

Components of Variation Due to Individual Order
(Partialling out the effects of both initial score and M.A.)

Source of Variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Adjusted Σy^2	Reduced d.f.	S.S.	M.S.	F.	Hypothesis*
Linear	1	289	220	992	252	536	467	59.20	1	33.12	33.12	25.48	Rejected
Quadratic	1	3	0	30	1	9	4	1.32	1	3.36	3.36	2.58	Accepted
Individual Order	2	292	220	1022	253	545	471	60.52					

* The hypothesis tested is a null hypothesis concerning the variation in the same row by using residual variance in Table 8 as a basis. For instance, the hypothesis concerning the quadratic component is that there is no significant quadratic relationship between final scores and individual orders when the effects of both initial score and mental ages have been removed.

TABLE 14
Summary of Results for Variation Due to Individual Order

Individual		Y	X_1	X_2	$\xi'_1(i)$	$\xi'_2(i)$
First	Sums of Scores	404	358	878	+1	+1
	Mean	22.4445	19.8889	48.7777		
Second	Sums of Scores	362	317	812	0	−2
	Mean	20.1111	17.6111	45.1111		
Third	Sums of Scores	302	269	689	−1	+1
	Mean	16.7778	14.9445	38.2777		

3. When final score results alone were considered the following components were found to be significant:

Main effects	*Interactions*
Sex	Sex × Grade
Grade	Grade × Scholastic Standing (doubtful)
Scholastic standing	
Individual order	

4. After the effect of initial score and mental age was partialled out, all the interactions were insignificant and there was no significant difference among grades. The only factors which remained significant were:

   Sex (doubtful)
   Scholastic standing
   Individual order

5. Regression equations were developed to estimate the observed and the predicted final scores from classifications of scholastic standings and from individual orders of educational development. The method was used of fitting by orthogonal polynomials. The analysis of variance was used to test the goodness of graduation.

The methods and procedures illustrated in this investigation are capable of extensive application. The conventional methods for attacking problems of this kind are usually inexact, of low precision, and incapable of yielding conclusive results. The introduction of covariance methods into factorial design will likely prove to be very useful, if not indispensable, in many psychological and educational problem.

It is customary to consider variation within sub-groups as residual, or experimental error. The results from this study point to the

fact that individual differences within a group are so strikingly significant that we cannot regard them as a residual. Where experimental designs take this factor into account, increase in precision may be expected, as has been convincingly demonstrated in this investigation.

The mathematical formulation and the solution underlying the method and analysis applied in Part I follow.

PART II

The derivation of sums of squares for each source of variation have been shown in the previous paper.* It is not necessary that it be repeated here. When the analysis of covariance technique is used, we have to calculate the sum of products. Let us consider only the simplest way of calculating sum of products. Suppose we define

$X_i =$ the raw score of the i-th individual on X variable
$Y_i =$ the raw score of the i-th individual on Y variable
$x_i =$ the deviation from the mean $\overline{X}$ for the i-th individual
$y_i =$ the deviation from the mean $\overline{Y}$ for the i-th individual
$i = 1, 2, \cdots, N$
$N =$ the number of cases

We have no difficulty in getting

$$\begin{aligned} \sum_i x_i^2 &= \sum_i X_i^2 - \frac{(\sum_i X_i)^2}{N} \\ \sum_i y_i^2 &= \sum_i Y_i^2 - \frac{(\sum_i Y_i)^2}{N} \\ \sum_i x_i y_i &= \sum_i X_i Y_i - \frac{(\sum_i X_i)(\sum_i Y_i)}{N} \end{aligned} \tag{1}$$

It is customary that the second term of the right-hand member of each equation mentioned above is called the "correction" term. On the other hand, the first term is only the straightforward calculation of the sum of squares (or products) of all the raw scores. If we always keep in mind what the difference is between the "correction" term for the sum of squares and that for the sum of products, then we can get all the solutions.

* Johnson and Tsao, *op. cit.*, pp. 128-143.

As for our problem, we define

Y_{sijt} = the final standard score of the t-th individual of the j-th scholastic standing in the i-th grade and the s-th sex,

$X_{1_{sijt}}$ = the initial standard score of the t-th individual of the j-th scholastic standing in the i-th grade and the s-th sex,

$X_{2_{sijt}}$ = the mental age score of the t-th individual of the j-th scholastic standing in the i-th grade and the s-th sex,

where $s = 1, 2$; $i = 1, 2, 3$; $j = 1, 2, 3$; $t = 1, 2, 3$.

Assume that

$$a_1 = \sum_s \sum_i \sum_j \sum_t Y^2_{sijt}; \quad a_2 = \sum_s \sum_i \sum_j \sum_t X^2_{1_{sijt}}; \quad a_3 = \sum_s \sum_i \sum_j \sum_t X^2_{2_{sijt}};$$

$$a_4 = \sum_s \sum_i \sum_j \sum_t (X_{1_{sijt}} Y_{sijt}); \qquad a_5 = \sum_s \sum_i \sum_j \sum_t (X_{2_{sijt}} Y_{sijt});$$

$$a_6 = \sum_s \sum_i \sum_j \sum_t (X_{1_{sijt}} X_{2_{sijt}});$$

$$b_{11} = \sum_s \sum_i \sum_j \left\{ \frac{(\sum_t Y_{sijt})^2}{3} \right\}, \qquad b_{12} = \sum_s \sum_i \sum_t \left\{ \frac{(\sum_j Y_{sijt})^2}{3} \right\},$$

$$b_{13} = \sum_s \sum_j \sum_t \left\{ \frac{(\sum_i Y_{sijt})^2}{3} \right\}, \qquad b_{14} = \sum_i \sum_j \sum_t \left\{ \frac{(\sum_s Y_{sijt})^2}{2} \right\};$$

$$b_{21} = \sum_s \sum_i \sum_j \left\{ \frac{(\sum_t X_{1_{sijt}})^2}{3} \right\}, \qquad b_{22} = \sum_s \sum_i \sum_t \left\{ \frac{(\sum_j X_{1_{sijt}})^2}{3} \right\},$$

$$b_{23} = \sum_s \sum_j \sum_t \left\{ \frac{(\sum_i X_{1_{sijt}})^2}{3} \right\}, \qquad b_{24} = \sum_i \sum_j \sum_t \left\{ \frac{(\sum_s X_{1_{sijt}})^2}{2} \right\};$$

$$b_{31} = \sum_s \sum_i \sum_j \left\{ \frac{(\sum_t X_{2_{sijt}})^2}{3} \right\}, \qquad b_{32} = \sum_s \sum_i \sum_t \left\{ \frac{(\sum_i X_{2_{sijt}})^2}{3} \right\},$$

$$b_{33} = \sum_s \sum_j \sum_t \left\{ \frac{(\sum_i X_{2_{sijt}})^2}{3} \right\}, \qquad b_{34} = \sum_i \sum_j \sum_t \left\{ \frac{(\sum_s X_{2_{sijt}})^2}{2} \right\};$$

$$b_{41}=\sum_s\sum_i\sum_j\left\{\frac{(\sum_t X_{1sijt})(\sum_t Y_{sijt})}{3}\right\},\quad b_{42}=\sum_s\sum_i\sum_t\left\{\frac{(\sum_j X_{1sijt})(\sum_j Y_{sijt})}{3}\right\},$$

$$b_{43}=\sum_s\sum_j\sum_t\left\{\frac{(\sum_i X_{1sijt})(\sum_i Y_{sijt})}{3}\right\},\quad b_{44}=\sum_i\sum_j\sum_t\left\{\frac{(\sum_s X_{1sijt})(\sum_s Y_{sijt})}{2}\right\};$$

$$b_{51}=\sum_s\sum_i\sum_j\left\{\frac{(\sum_t X_{2sijt})(\sum_t Y_{sijt})}{3}\right\},\quad b_{52}=\sum_s\sum_i\sum_t\left\{\frac{(\sum_j X_{2sijt})(\sum_j Y_{sijt})}{3}\right\},$$

$$b_{53}=\sum_s\sum_j\sum_t\left\{\frac{(\sum_i X_{2sijt})(\sum_i Y_{sijt})}{3}\right\},\quad b_{54}=\sum_s\sum_j\sum_t\left\{\frac{(\sum_s X_{2sijt})(\sum_s Y_{sijt})}{2}\right\};$$

$$b_{61}=\sum_s\sum_i\sum_j\left\{\frac{(\sum_t X_{1sijt})(\sum_t X_{2sijt})}{3}\right\},\quad b_{62}=\sum_s\sum_i\sum_t\left\{\frac{(\sum_j X_{1sijt})(\sum_j X_{2sijt})}{3}\right\},$$

$$b_{63}=\sum_s\sum_j\sum_t\left\{\frac{(\sum_i X_{1sijt})(\sum_i X_{2sijt})}{3}\right\},\quad b_{64}=\sum_s\sum_j\sum_t\left\{\frac{(\sum_s X_{1sijt})(\sum_s X_{2sijt})}{2}\right\};$$

$$c_{11}=\sum_s\sum_i\left\{\frac{(\sum_j\sum_t Y_{sijt})^2}{9}\right\},\quad c_{12}=\sum_s\sum_j\left\{\frac{(\sum_i\sum_t Y_{sijt})^2}{9}\right\},\quad c_{13}=\sum_s\sum_t\left\{\frac{(\sum_i\sum_j Y_{sijt})^2}{9}\right\},$$

$$c_{14}=\sum_i\sum_j\left\{\frac{(\sum_s\sum_t Y_{sijt})^2}{6}\right\},\quad c_{15}=\sum_i\sum_t\left\{\frac{(\sum_s\sum_j Y_{sijt})^2}{6}\right\},\quad c_{16}=\sum_j\sum_t\left\{\frac{(\sum_s\sum_i Y_{sijt})^2}{6}\right\};$$

$$c_{21}=\sum_s\sum_i\left\{\frac{(\sum_j\sum_t X_{1sijt})^2}{9}\right\},\quad c_{22}=\sum_s\sum_j\left\{\frac{(\sum_i\sum_t X_{1sijt})^2}{9}\right\},\quad c_{23}=\sum_s\sum_t\left\{\frac{(\sum_i\sum_j X_{1sijt})^2}{9}\right\},$$

$$c_{24}=\sum_i\sum_j\left\{\frac{(\sum_s\sum_t X_{1sijt})^2}{6}\right\},\quad c_{25}=\sum_i\sum_t\left\{\frac{(\sum_s\sum_j X_{1sijt})^2}{6}\right\},\quad c_{26}=\sum_j\sum_t\left\{\frac{(\sum_s\sum_i X_{1sijt})^2}{6}\right\};$$

$$c_{31}=\sum_s\sum_i\left\{\frac{(\sum_j\sum_t X_{2sijt})^2}{9}\right\},\quad c_{32}=\sum_s\sum_j\left\{\frac{(\sum_i\sum_t X_{2sijt})^2}{9}\right\},\quad c_{33}=\sum_s\sum_t\left\{\frac{(\sum_i\sum_j X_{2sijt})^2}{9}\right\},$$

$$c_{34}=\sum_i\sum_j\left\{\frac{(\sum_s\sum_t X_{2sijt})^2}{6}\right\},\quad c_{35}=\sum_i\sum_t\left\{\frac{(\sum_s\sum_j X_{2sijt})^2}{6}\right\},\quad c_{36}=\sum_j\sum_t\left\{\frac{(\sum_s\sum_i X_{2sijt})^2}{6}\right\};$$

$$c_{41}=\sum_s\sum_i\left\{\frac{\left(\sum_s\sum_t X_{1sijt}\right)\left(\sum_j\sum_t Y_{sijt}\right)}{9}\right\},\quad c_{42}=\sum_s\sum_j\left\{\frac{\left(\sum_i\sum_t X_{1sijt}\right)\left(\sum_i\sum_t Y_{sijt}\right)}{9}\right\},$$

$$c_{43}=\sum_s\sum_t\left\{\frac{\left(\sum_i\sum_j X_{1sijt}\right)\left(\sum_i\sum_j Y_{sijt}\right)}{9}\right\},\quad c_{44}=\sum_i\sum_j\left\{\frac{\left(\sum_s\sum_t X_{1sijt}\right)\left(\sum_s\sum_t Y_{sijt}\right)}{6}\right\};$$

$$c_{45}=\sum_i\sum_t\left\{\frac{\left(\sum_s\sum_j X_{1sijt}\right)\left(\sum_s\sum_j Y_{sijt}\right)}{6}\right\},\quad c_{46}=\sum_j\sum_t\left\{\frac{\left(\sum_s\sum_i X_{1sijt}\right)\left(\sum_s\sum_i Y_{sijt}\right)}{6}\right\};$$

$$c_{51}=\sum_s\sum_i\left\{\frac{\left(\sum_j\sum_t X_{2sijt}\right)\left(\sum_j\sum_t Y_{sijt}\right)}{9}\right\},\quad c_{52}=\sum_s\sum_j\left\{\frac{\left(\sum_i\sum_t X_{2sijt}\right)\left(\sum_i\sum_t Y_{sijt}\right)}{9}\right\},$$

$$c_{53}=\sum_s\sum_t\left\{\frac{\left(\sum_i\sum_j X_{2sijt}\right)\left(\sum_i\sum_j Y_{sijt}\right)}{9}\right\},\quad c_{54}=\sum_i\sum_j\left\{\frac{\left(\sum_s\sum_t X_{2sijt}\right)\left(\sum_s\sum_t Y_{sijt}\right)}{6}\right\},$$

$$c_{55}=\sum_i\sum_t\left\{\frac{\left(\sum_s\sum_j X_{2sijt}\right)\left(\sum_s\sum_j Y_{sijt}\right)}{6}\right\},\quad c_{56}=\sum_j\sum_t\left\{\frac{\left(\sum_s\sum_i X_{2sijt}\right)\left(\sum_s\sum_i Y_{sijt}\right)}{6}\right\};$$

$$c_{61}=\sum_s\sum_i\left\{\frac{\left(\sum_j\sum_t X_{1sijt}\right)\left(\sum_j\sum_t X_{2sijt}\right)}{9}\right\},\quad c_{62}=\sum_s\sum_j\left\{\frac{\left(\sum_i\sum_t X_{1sijt}\right)\left(\sum_i\sum_t X_{2sijt}\right)}{9}\right\},$$

$$c_{63}=\sum_s\sum_t\left\{\frac{\left(\sum_i\sum_j X_{1sijt}\right)\left(\sum_i\sum_j X_{2sijt}\right)}{9}\right\},\quad c_{64}=\sum_i\sum_j\left\{\frac{\left(\sum_s\sum_t X_{1sijt}\right)\left(\sum_s\sum_t X_{2sijt}\right)}{6}\right\},$$

$$c_{65}=\sum_i\sum_t\left\{\frac{\left(\sum_s\sum_j X_{1sijt}\right)\left(\sum_s\sum_j X_{2sijt}\right)}{6}\right\},\quad c_{66}=\sum_j\sum_t\left\{\frac{\left(\sum_s\sum_i X_{1sijt}\right)\left(\sum_s\sum_i X_{2sijt}\right)}{6}\right\};$$

$$d_{11}=\sum_s\left\{\frac{\left(\sum_i\sum_j\sum_t Y_{sijt}\right)^2}{27}\right\},\quad d_{12}=\sum_i\left\{\frac{\left(\sum_s\sum_j\sum_t Y_{sijt}\right)^2}{18}\right\},\quad d_{13}=\sum_j\left\{\frac{\left(\sum_s\sum_i\sum_t Y_{sijt}\right)^2}{18}\right\},$$

$$d_{14}=\sum_t\left\{\frac{\left(\sum_s\sum_i\sum_j Y_{sijt}\right)^2}{18}\right\};$$

$$d_{21}=\sum_s\left\{\frac{\left(\sum_i\sum_j\sum_t X_{1sijt}\right)^2}{27}\right\},\quad d_{22}=\sum_i\left\{\frac{\left(\sum_s\sum_j\sum_t X_{1sijt}\right)^2}{18}\right\},\quad d_{23}=\sum_j\left\{\frac{\left(\sum_s\sum_i\sum_t X_{1sijt}\right)^2}{18}\right\},$$

$$d_{24}=\sum_{t}\left\{\frac{(\sum_{s}\sum_{i}\sum_{j}X_{1sijt})^2}{18}\right\};$$

$$d_{31}=\sum_{s}\left\{\frac{(\sum_{i}\sum_{j}\sum_{t}X_{2sijt})^2}{27}\right\},\qquad d_{32}=\sum_{i}\left\{\frac{(\sum_{s}\sum_{j}\sum_{t}X_{2sijt})^2}{18}\right\},\qquad d_{33}=\sum_{j}\left\{\frac{(\sum_{s}\sum_{i}\sum_{t}X_{2sijt})^2}{18}\right\},$$

$$d_{34}=\sum_{t}\left\{\frac{(\sum_{s}\sum_{i}\sum_{j}X_{2sijt})^2}{18}\right\};$$

$$d_{41}=\sum_{s}\left\{\frac{(\sum_{i}\sum_{j}\sum_{t}X_{1sijt})(\sum_{i}\sum_{j}\sum_{t}Y_{sijt})}{27}\right\},\qquad d_{42}=\sum_{i}\left\{\frac{(\sum_{s}\sum_{j}\sum_{t}X_{1sijt})(\sum_{s}\sum_{j}\sum_{t}Y_{sijt})}{18}\right\},$$

$$d_{43}=\sum_{j}\left\{\frac{(\sum_{s}\sum_{i}\sum_{t}X_{1sijt})(\sum_{s}\sum_{i}\sum_{t}Y_{sijt})}{18}\right\},\qquad d_{44}=\sum_{t}\left\{\frac{(\sum_{s}\sum_{i}\sum_{j}X_{1sijt})(\sum_{s}\sum_{i}\sum_{j}Y_{sijt})}{18}\right\};$$

$$d_{51}=\sum_{s}\left\{\frac{(\sum_{i}\sum_{j}\sum_{t}X_{2sijt})(\sum_{i}\sum_{j}\sum_{t}Y_{sijt})}{27}\right\},\qquad d_{52}=\sum_{i}\left\{\frac{(\sum_{s}\sum_{j}\sum_{t}X_{2sijt})(\sum_{s}\sum_{j}\sum_{t}Y_{sijt})}{18}\right\},$$

$$d_{53}=\sum_{j}\left\{\frac{(\sum_{s}\sum_{i}\sum_{t}X_{2sijt})(\sum_{s}\sum_{i}\sum_{t}Y_{sijt})}{18}\right\},\qquad d_{54}=\sum_{t}\left\{\frac{(\sum_{s}\sum_{i}\sum_{j}X_{2sijt})(\sum_{s}\sum_{i}\sum_{j}Y_{sijt})}{18}\right\};$$

$$d_{61}=\sum_{s}\left\{\frac{(\sum_{i}\sum_{j}\sum_{t}X_{1sijt})(\sum_{i}\sum_{j}\sum_{t}X_{2sijt})}{27}\right\},\qquad d_{62}=\sum_{i}\left\{\frac{(\sum_{s}\sum_{j}\sum_{t}X_{1sijt})(\sum_{s}\sum_{j}\sum_{t}X_{2sijt})}{18}\right\},$$

$$d_{63}=\sum_{j}\left\{\frac{(\sum_{s}\sum_{i}\sum_{t}X_{1sijt})(\sum_{s}\sum_{i}\sum_{t}X_{2sijt})}{18}\right\},\qquad d_{64}=\sum_{t}\left\{\frac{(\sum_{s}\sum_{i}\sum_{j}X_{1sijt})(\sum_{s}\sum_{i}\sum_{j}X_{2sijt})}{18}\right\};$$

$$e_1=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{sijt})^2}{54};\qquad e_2=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{1sijt})^2}{54};\qquad e_3=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{2sijt})^2}{54};$$

$$e_4=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{1sijt})(\sum_{s}\sum_{i}\sum_{j}\sum_{t}Y_{sijt})}{54};\qquad e_5=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{2sijt})(\sum_{s}\sum_{i}\sum_{j}\sum_{t}Y_{sijt})}{54};$$

$$e_6=\frac{(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{1sijt})(\sum_{s}\sum_{i}\sum_{j}\sum_{t}X_{2sijt})}{54}.$$

Then all the calculations of the sum of squares and sum of products for each source of variation will be summarized in Table 1.

TABLE 1

Sum of Squares and of Products for Each Source of Variation

Source of Variation	D.F.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$
Sex × Grade × Schol. × Individual	8	$a_1 - \sum_{k=1}^{4} b_{1k} + \sum_{k=1}^{6} c_{1k} - \sum_{k=1}^{4} d_{1k} + e_1$	$a_2 - \sum_{k=1}^{4} b_{2k} + \sum_{k=1}^{6} c_{2k} - \sum_{k=1}^{4} d_{2k} + e_2$	$a_3 - \sum_{k=1}^{4} b_{3k} + \sum_{k=1}^{6} c_{3k} - \sum_{k=1}^{4} d_{3k} + e_3$	$a_4 - \sum_{k=1}^{4} b_{4k} + \sum_{k=1}^{6} c_{4k} - \sum_{k=1}^{4} d_{4k} + e_4$	$a_5 - \sum_{k=1}^{4} b_{5k} + \sum_{k=1}^{6} c_{5k} - \sum_{k=1}^{4} d_{5k} + e_5$	$a_6 - \sum_{k=1}^{4} b_{6k} + \sum_{k=1}^{6} c_{6k} - \sum_{k=1}^{4} d_{6k} + e_6$
Sex × Grade × Schol.	4	$b_{11} - (c_{11} + c_{12} + c_{14}) + (d_{11} + d_{12} + d_{13}) - e_1$	$b_{21} - (c_{21} + c_{22} + c_{24}) + (d_{21} + d_{22} + d_{23}) - e_2$	$b_{31} - (c_{31} + c_{32} + c_{34}) + (d_{31} + d_{32} + d_{33}) - e_3$	$b_{41} - (c_{41} + c_{42} + c_{44}) + (d_{41} + d_{42} + d_{43}) - e_4$	$b_{51} - (c_{51} + c_{52} + c_{52}) + (d_{51} + d_{52} + d_{53}) - e_5$	$b_{61} - (c_{61} + c_{62} + c_{64}) + (d_{61} + d_{62} + d_{63}) - e_6$
Sex × Grade × Individual	4	$b_{12} - (c_{11} + c_{13} + c_{15}) + (d_{11} + d_{12} + d_{14}) - e_1$	$b_{22} - (c_{21} + c_{23} + c_{25}) + (d_{21} + d_{22} + d_{24}) - e_2$	$b_{32} - (c_{31} + c_{33} + c_{35}) + (d_{31} + d_{32} + d_{34}) - e_3$	$b_{42} - (c_{41} + c_{43} + c_{45}) + (d_{41} + d_{42} + d_{44}) - e_4$	$b_{52} - (c_{51} + c_{53} + c_{55}) + (d_{51} + d_{52} + d_{54}) - e_5$	$b_{62} - (c_{61} + c_{63} + c_{65}) + (d_{61} + d_{62} + d_{64}) - e_6$
Sex × Schol. × Individual	4	$b_{13} - (c_{12} + c_{13} + c_{16}) + (d_{11} + d_{13} + d_{14}) - e_1$	$b_{23} - (c_{22} + c_{23} + c_{26}) + (d_{21} + d_{23} + d_{24}) - e_2$	$b_{33} - (c_{32} + c_{33} + c_{36}) + (d_{31} + d_{33} + d_{34}) - e_3$	$b_{43} - (c_{42} + c_{43} + c_{46}) + (d_{41} + d_{43} + d_{44}) - e_4$	$b_{53} - (c_{52} + c_{53} + c_{56}) + (d_{51} + d_{53} + d_{54}) - e_5$	$b_{63} - (c_{62} + c_{63} + c_{66}) + (d_{61} + d_{63} + d_{64}) - e_6$
Grade × Schol. × Individual	8	$b_{14} - (c_{14} + c_{15} + c_{16}) + (d_{12} + d_{13} + d_{14}) - e_1$	$b_{24} - (c_{24} + c_{25} + c_{26}) + (d_{22} + d_{23} + d_{24}) - e_2$	$b_{34} - (c_{34} + c_{35} + c_{36}) + (d_{32} + d_{33} + d_{34}) - e_3$	$b_{44} - (c_{44} + c_{45} + c_{46}) + (d_{42} + d_{43} + d_{44}) - e_4$	$b_{54} - (c_{54} + c_{55} + c_{56}) + (d_{52} + d_{53} + d_{54}) - e_5$	$b_{64} - (c_{64} + c_{65} + c_{66}) - (d_{62} + d_{63} + d_{64}) - e_6$

TABLE 1 (Cont.)
Sum of Squares and of Products for Each Source of Variation

Source of Variation	D.F.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$
Sex × Grade	2	$c_{11} - (d_{11} + d_{12}) + e_1$	$c_{21} - (d_{21} + d_{22}) + e_2$	$c_{31} - (d_{31} + d_{32}) + e_3$	$c_{41} - (d_{41} + d_{42}) + e_4$	$c_{51} - (d_{51} + d_{52}) + e_5$	$c_{61} - (d_{61} + d_{62}) + e_6$
Sex × Schol.	2	$c_{12} - (d_{11} + d_{13}) + e_1$	$c_{22} - (d_{21} + d_{23}) + e_2$	$c_{32} - (d_{31} + d_{33}) + e_3$	$c_{42} - (d_{41} + d_{43}) + e_4$	$c_{52} - (d_{51} + d_{53}) + e_5$	$c_{62} - (d_{61} + d_{63}) + e_6$
Sex × Individual	2	$c_{13} - (d_{11} + d_{14}) + e_1$	$c_{23} - (d_{21} + d_{24}) + e_2$	$c_{33} - (d_{31} + d_{34}) + e_3$	$c_{43} - (d_{41} + d_{44}) + e_4$	$c_{53} - (d_{51} + d_{54}) + e_5$	$c_{63} - (d_{61} + d_{64}) + e_6$
Grade × Schol.	4	$c_{14} - (d_{12} + d_{13}) + e_1$	$c_{24} - (d_{22} + d_{23}) + e_2$	$c_{34} - (d_{32} + d_{33}) + e_3$	$c_{44} - (d_{42} + d_{43}) + e_4$	$c_{54} - (d_{52} + d_{53}) + e_5$	$c_{64} - (d_{62} + d_{63}) + e_6$
Grade × Individual	4	$c_{15} - (d_{12} + d_{14}) + e_1$	$c_{25} - (d_{22} + d_{24}) + e_2$	$c_{35} - (d_{32} - d_{34}) + e_3$	$c_{45} - (d_{42} + d_{44}) + e_4$	$c_{55} - (d_{52} + d_{54}) + e_5$	$c_{65} - (d_{62} + d_{64}) + e_6$
Schol. × Individual	4	$c_{16} - (d_{13} + d_{14}) + e_1$	$c_{26} - (d_{23} + d_{24}) + e_2$	$c_{36} - (d_{33} + d_{34}) + e_3$	$c_{46} - (d_{43} + d_{44}) + e_4$	$c_{56} - (d_{53} + d_{54}) + e_5$	$c_{66} - (d_{63} + d_{64}) + e_6$
Sex	1	$d_{11} - e_1$	$d_{21} - e_2$	$d_{31} - e_3$	$d_{41} - e_4$	$d_{51} - e_5$	$d_{61} - e_6$
Grade	2	$d_{12} - e_1$	$d_{22} - e_2$	$d_{32} - e_3$	$d_{42} - e_4$	$d_{52} - e_5$	$d_{62} - e_6$
Scholastic	2	$d_{13} - e_1$	$d_{23} - e_2$	$d_{33} - e_3$	$d_{43} - e_4$	$d_{53} - e_5$	$d_{63} - e_6$
Individual	2	$d_{14} - e_1$	$d_{24} - e_2$	$d_{34} - e_3$	$d_{44} - e_4$	$d_{54} - e_5$	$d_{64} - e_6$
Total	53	$a_1 - e_1$	$a_2 - e_2$	$a_3 - e_3$	$a_4 - e_4$	$a_5 - e_5$	$a_6 - e_6$

Now we consider how to eliminate the effect of initial scores in our study. By using regression of y on x, we have

$$\tilde{y} = bx_1, \tag{2}$$

where $\tilde{y}$ denotes the predicted deviation from the mean of final scores, x_1 denotes the deviation from the mean of initial scores, and b is regression coefficient.

Minimizing the difference between the actual y score and the predicted y score, we have

$$\chi^2 = \Sigma(y - bx_1)^2. \tag{3}$$

Differentiating χ^2 with respect to b, we obtain

$$b = \frac{\Sigma xy}{\Sigma x^2}, \tag{4}$$

Consequently, we get

$$\chi^2 = \Sigma(y - bx_1)^2 = \Sigma y^2 - 2b\Sigma x_1 y + b^2 \Sigma x_1^2. \tag{5}$$

So the coefficients for adjustying Σy^2 for all sources of variation are as follows:

Σy^2	$\Sigma x_1 y$	Σx_1^2
1	$-2b$	b^2

Remembering that the value of b comes from the error term (when we make the test of significance of interactions) or the residual (when we make a complete analysis), let us consider this term only. By simple algebraic manipulation, we have

$$\begin{aligned} \chi^2 &= \Sigma y^2 - 2b\Sigma x_1 y + b^2 \Sigma x_1^2 \\ &= \Sigma y^2 - b\Sigma x_1 y = \Sigma y_2 - \frac{(\Sigma x_1 y)^2}{\Sigma x_1^2}. \end{aligned} \tag{6}$$

Therefore the degrees of freedom in error or residual will be diminished by 1 because of the restriction of $b\Sigma x_1 y$ or $\frac{(\Sigma x_1 y)^2}{\Sigma x_1^2}$ in making adjustment.

The value of b used for adjusting Σy^2 for all sources of variation is a statistical estimate subject to errors of random sampling. So it is advisable in making an exact test of significance that we use the reduced value* instead of the adjusted value. For instance, in Table

* Fisher, R. A. Statistical methods for research workers, (7th edition). London: Oliver and Boyd, 1938, pp. 281-294.

6 of Part I, the reduced value of the sum of squares for the variation of sex is obtained by the following procedure:

Taking those parts of Table 6 of Part I which refer to residual and sex only, we obtain the reduced values of Σy^2, respectively, for the residual and for the total, by deducting from Σy^2 in each case the quantity

$$\frac{(\Sigma xy)^2}{\Sigma x^2},$$

derived from the same line. This gives for the residual the reduced value of Σy^2, 70.86, as in Table 6 of Part I. For the total we have

$$216 - 138.75 = 77.25,$$

corresponding to 46 degrees of freedom. Subtracting the reduced Σy^2 for residual from the reduced Σy^2 for total, we find the reduced Σy^2 ascribable to the 1 degree of freedom for sex to be 6.39, which is the value to be compared with the reduced Σy^2 for residual, in making an exact test. The whole procedure is shown in Table 2.

TABLE 2
Illustration of Test of Significance with Reduced Σy^2
(Partialling out the effect of X_1)

Source of Variation	d.f.	Σy^2	$\Sigma x_1 y$	Σx_1^2	Reduced d.f.	Reduced S.S.	Reduced M.S.	Reduced F.	Reduced Hypothesis
Residual	46	186	179	278	45	70.86	1.57	—	—
Sex	1	30	24	19	1	6.39	6.39	4.07	Remain in doubt
Total	47	216	203	297	46	77.25			

Finally, we work out the regression coefficients for adjusting Σy^2 in partialling out both the effects of initial scores and M.A. By using the regression of y on x_1 and x_2, we have

$$\tilde{y} = b_1 x_1 + b_2 x_2,$$

where $\tilde{y}$ denotes the predicted deviation final score, x_1 denotes the deviation initial score, x_2 denotes the deviation M. A., b_1 is the partial regression coefficient of x_1, and b_2 is the partial regression coefficient of x_2.

Minimizing the difference between the actual y score and the predicted y score, we have

$$\chi^2 = \Sigma (y - b_1 x_1 - b_2 x_2)^2. \tag{7}$$

Differentiating χ^2 with respect to b_1 and b_2, we have

$$\begin{aligned} b_1 \sum x_1^2 + b_2 \sum x_1 x_2 &= \sum x_1 y \\ b_1 \sum x_1 x_2 + b_2 \sum x_2^2 &= \sum x_2 y \,. \end{aligned} \tag{8}$$

By using the determinant method, we have

$$\begin{aligned} b_1 &= \frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \\ b_2 &= \frac{\sum x_1^2 \sum x_2 y - \sum x_1 x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \,. \end{aligned} \tag{9}$$

Consequently, we get

$$\begin{aligned} \chi^2 = \sum (y - b_1 x_1 - b_2 x_2)^2 &= \sum y^2 + b_1^2 \sum x_1^2 \\ &+ b_2^2 \sum x_2^2 - 2b_1 \sum x_1 y - 2b_2 \sum x_2 y + 2b_1 b_2 \sum x_1 x_2 \,. \end{aligned} \tag{10}$$

So the coefficients for adjusting $\sum y$ for all sources of variation are as follows:

$\sum y^2$	$\sum x_1^2$	$\sum x_2^2$	$\sum x_1 y$	$\sum x_2 y$	$\sum x_1 x_2$
1	b_1^2	b_2^2	$-2b_1$	$-2b_2$	$2b_1 b_2$.

The values of b_1 and b_2 come from the error term or the residual term. Let us consider this term only; we have, by troublesome algebraic manipulation,

$$\begin{aligned} \chi^2 &= \sum y^2 + b_1^2 \sum x_1^2 + b_2^2 \sum x_2^2 - 2b_1 \sum x_1 y \\ &\quad - 2b_2 \sum x_2 y + 2b_1 b_2 \sum x_1 x_2 \\ &= \sum y^2 - b_1 \sum x_1 y - b_2 \sum x_2 y \,. \end{aligned} \tag{11}$$

Therefore, the degrees of freedom in error or residual will be diminished by 2 because of the two restrictions of $b_1 \sum x_1 y$ and $b_2 \sum x_2 y$ in making adjustment.

Once more in making an exact test of significance for each source of variation we use the reduced $\sum y^2$ instead of the adjusted $\sum y^2$. The calculation of the reduced $\sum y^2$ is similar to the one used earlier. To give an example, in Table 8 of Part I, the reduced $\sum y^2$ for the variation of grade is obtained by the following process:

Taking those parts of Table 8 of Part I which refer to residual and grade only, we obtain the reduced $\sum y^2$, respectively, for the residual and for the total, by deducting from $\sum y^2$ in each case the quantity

$$\frac{\sum x_2^2 (\sum x_1 y)^2 - \sum x_1 y \sum x_2 y \sum x_1 x_2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} + \frac{\sum x_1^2 (\sum x_2 y)^2 - \sum x_1 y \sum x_2 y \sum x_1 x_2}{\sum x_1^2 \sum x_2^2 (\sum x_1 x_2)^2}$$

or

$$\frac{\sum x_2^2(\sum x_1 y)^2 - 2\sum x_1 y \sum x_2 y \sum x_1 x_2 + \sum x_1^2(\sum x_2 y)^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

derived from the same line. This gives the reduced $\sum y^2$ for residual, 57.21, as in Table 8 of Part I. For the total we have

$$248 - 87.90 = 60.10\,,$$

corresponding to 46 degrees of freedom. Subtracting 57.21 from 60.10, we find the reduced $\sum y^2$ ascribable to the 2 degrees of freedom for grade to be 2.89. The whole procedure of making an exact test of significance is summarized in Table 3:

TABLE 3
Illustrations of Test of Significance with Reduced Σy^2
(Partialling out the effects of X_1 and X_2)

Source of Variation	d.f.	Σy^2	Σx_1^2	Σx_2^2	$\Sigma x_1 y$	$\Sigma x_2 y$	$\Sigma x_1 x_2$	Reduced d.f.	Reduced S.S.	Reduced M.S.	Reduced F.	Reduced Hypothesis
Residual	46	186	278	6675	179	697	678	44	57.21	1.30	—	—
Grade	2	62	83	5837	70	594	697	2	2.89	1.45	1.12	Accepted
Total	48	248	261	12512	249	1291	1375	46	60.10			

ERRATUM

Dr. Frederick B. Davis points out that the minus sign should have been a plus sign in formula (6), p. 59, of his article, "The Reliability of Component Scores," in Volume 10, No. 1 of this journal.